

Mon-Khmer Prosodic Morphology : System-Independent and System-Dependent Motivations

著者	宮腰 幸一
journal or publication title	論叢現代文化・公共政策
volume	2
page range	1-60
year	2005-10-31
URL	http://hdl.handle.net/2241/17974

Mon-Khmer Prosodic Morphology: System-Independent and System-Dependent Motivations*

Koichi Miyakoshi

This paper addresses some issues of prosodic morphology and linguistic theory, proposing an Optimality-Theoretic analysis of Mon-Khmer prosody and reduplication. Mon-Khmer languages have apparently unusual types of reduplication in terms of the base-reduplicant association pattern and the reduplicant's prosodic form. This paper argues that they can nevertheless be explained as instances of the emergence of the unmarked, in the sense that the independently-motivated canonical structure of a language is utilized for a new function and makes its appearance. It further argues that, generally, what reduplicated form a given language has depends crucially on what prosodic structure it has, adducing evidence from various types of reduplication in the world's languages that exhibit such prosodic dependency. The paper also explores motivations for such emergence of the unmarked and linguistic dependency, suggesting two types of general constraints on language, which are formulated in terms of system-(in)dependency.

1. Introduction

A language is a system of systems with interdependent elements. What element a given language system has depends strongly on what other elements it has, where the system is of any type: phonological, semantic or symbolic. This strong dependency among elements is one of the essential properties of human language (or of systems in general for that matter; cf. von Bertalanffy 1968, Laszlo 1972). It is therefore exceedingly rare that a given language entity is completely independent of other entities of the language. Many linguists have pointed out this interdependent/interconnected nature of linguistic entities, making claims like the following: "nothing in language stands by itself but each of its elements acts as a part of a whole" (Jakobson and Waugh 1979: 166; see also de Saussure 1916/1983: 50-51, Sapir 1921: 182, Jakobson 1962: 202, and the references cited below).

For example, consider the following consonantal inventories of hypothetical languages:¹

(1) Consonantal Inventories of Hypothetical Languages

a. Consonantal Inventory 1

p	*	*	t	*	*	k	*	*
b	*	*	d	*	*	g	*	*
m	*	*	n	*	*	ŋ	*	*

b. Consonantal Inventory 2

*	*	*	*	t̚	*	*	q	ʔ
b	*	ɗ	*	*	*	g	*	*
*	ɱ	*	n	*	ɲ	*	*	*

(Lindblom 1992: 137)

The consonantal inventory in (1a) appears to constitute a well-aligned system with interdependent elements, whereas that in (1b) appears to be just a random list of segments, which are mostly independent of one another. In fact, human languages strongly tend to have a highly (if not completely) systematic paradigm similar to (1a), rather than a random list like (1b), as exemplified by the following consonantal inventories of actual languages (cf. Sapir 1921: chapter 8; Trubetzkoy 1939/1969: chapter 3; Postal 1968: chapters 4 and 8; Lindblom 1992; Clements 2003):

(2) Consonantal Inventories of Actual Languages

a. Consonantal (Nasal) Inventory of Sui

m	n	ɲ	ŋ
m̥	n̥	ɲ̥	ŋ̥
m̥̃	ñ̥	ɲ̥̃	ŋ̥̃

b. Consonantal (Affricate) Inventory of Chipewyan

t̪θ	ts	tʃ	tʃ̥
t̪θ ^h	ts ^h	tʃ ^h	tʃ̥ ^h
t̪θ'	ts'	tʃ'	tʃ̥'

c. Consonantal (Stop) Inventory of Alawa

b	d	ɗ	ɗ̥	g
m̥b	n̥d	ɲ̥ɗ	ɲ̥ɗ̥	ŋ̥g
m	n	ɲ	ɲ̥	ŋ

(Lindblom 1992: 137)

In the case of Sui (a Tai language), for example, it is apparent that the twelve nasals are dependent on one another in terms of the place and manner of articulation. Likewise for the set of affricates in Chipewyan (an Athapaskan language) and for the set of stops in Alawa (an Australian language). These consonantal inventories suggest that what phonological segment a given language has depends strongly on what other segments it has. It seems to be this strong dependency among a set of segments that their exquisite systematicity is attributed to (their systematicity and dependency will be quantified later).

The present study demonstrates that this strong dependency holds among prosodic structures as well, and suggests a few general constraints that ensure the emergence of such linguistic dependency. It argues for them, adducing evidence primarily from reduplication in Temiar and Semai, the two main Mon-Khmer (and/or Aslian) languages spoken in the Malay peninsula.²

Temiar and Semai have several types of reduplication, which can be categorized into

the following five types:^{3,4}

(3) Five types of reduplication in Mon-Khmer languages:

a. Type 1: C_1C_f prefixation to the heavy syllable of the base

Base: a biconsonantal root

Meaning: active voice and continuative aspect

$C_1V(:)C_f \rightarrow C_1C_f.C_1V(:)C_f$ ⁵

- | | | | |
|---------|------|------------------------|----------------------|
| Temiar: | i. | kɔ:w → kw .kɔ:w | ‘to be calling’ |
| | ii. | ca:ʔ → cʔ .ca:ʔ | ‘to be eating’ |
| | iii. | lug → lg .lug | ‘to be laughing’ |
| Semai: | iv. | koh → kh .koh | ‘to be chopping off’ |
| | v. | ci:p → cp .ci:p | ‘to be walking’ |
| | vi. | kʉ:ʔ → kʔ .kʉ:ʔ | ‘to be vomiting’ |

b. Type 2: C_f infixation immediately to the left of the heavy syllable of the base

Base: a bi-/tri-consonantal root (and the causative affix *r/tr*)⁶

Meaning: active (causative) voice and continuative aspect

$C_1(r/tr).C V(:)C_f \rightarrow C_1(/t.r)C_f.C V(:)C_f$

- | | | | |
|---------|------|-----------------------------|-------------------------|
| Temiar: | i. | s.lɔg → sg .lɔg | ‘to be lying down’ |
| | ii. | sr.lɔg → s.rg .lɔg | ‘to be laying down’ |
| | iii. | tr.kɔ:w → t.rw .kɔ:w | ‘to be calling’ |
| Semai: | iv. | c.ʔu:l → cl .ʔu:l | ‘to be choking’ |
| | v. | b.he:ʔ → bʔ .he:ʔ | ‘to be being satisfied’ |
| | vi. | c.ləh → ch .ləh | ‘to be going down’ |

c. Type 3: C_1 prefixation with the fixed vowel *a* to the heavy syllable of the base

Base: a biconsonantal root

Meaning: i. active voice and simulfactive aspect

ii. resultatives

$C_1V(:)C_f \rightarrow C_1a.C_1V(:)C_f$

- | | | | |
|---------|------|------------------------|-------------------------------|
| Temiar: | i. | kɔ:w → ka .kɔ:w | ‘to call’ |
| Semai: | ii. | rɔc → ra .rɔc | ‘to be uprooted (of a tuber)’ |
| | iii. | cɛ:s → ca .cɛ:s | ‘to be torn off’ |

d. Type 4: C_1C_f prefixation to the stem base

Base: a polysyllabic stem

Meaning: “expressive” (prolongation or continuous repetition in time)⁷

$C_1(.CC).C V(:)C_f \rightarrow C_1C_f.C_1(.CC).C V(:)C_f$

- | | | | |
|---------|------|----------------------------------|------------------------------------|
| Temiar: | i. | b.guy → by .b.guy | ‘to waft (smoke)’ |
| | ii. | r.we:g → rg .r.we:g | ‘to stand conspicuously upright’ |
| | iii. | k.rd.lã:d → kd .k.rd.lã:d | ‘curly hair’ |
| Semai: | iv. | d.ŋɔh → dh .d.ŋɔh | ‘appearance of nodding constantly’ |
| | v. | g.hup → gp .g.hup | ‘irritation on skin’ |
| | vi. | b.ʔəl → bl .b.ʔəl | ‘painful embarrassment’ |

e. Type 5: CV(:)C_r suffixation to the stem base⁸

Base: a polysyllabic stem

Meaning: “expressive” (repetition at intervals of time)

C₁(.CV).CV(:)C_r → C₁(.CV).CV(:)C_r.CV(:)C_r

Temiar:	i.	k.ra.hab → k.ra.hab. hab	‘lip-smacking’
	ii.	k.ra.lɔg → k.ra.lɔg. lɔg .	‘sound of heavy footsteps’
	iii.	c.ra.ʔu:k → c.ra.ʔu:k. ʔu:k	‘stomach queaziness’
Semai:	iv.	d.yõ:l → d.yõ:l.yõ:l	‘the appearance of an object floating down a river and getting stuck here and there’
	v.	k.na.rã:c → k.na.rã:c. rã:c	‘repeated pains of deep wound’
	vi.	g.ra.yul → g.ra.yul. yul	‘several people shaking something repeatedly’

As indicated earlier (note 4), these forms are not phonetically realized as such. A short vowel is epenthesized between two consecutive consonants. In Temiar, the short mid vowel *ɛ* is used in closed syllables, and the schwa *ə* in open syllables. For example, (3ai) and (3di) are phonetically realized as *kewkɔ:w* and *beybəguy*, respectively. What should be noted here is that, a few exceptions aside, these vowels consistently appear in the environments specified just above. They are therefore likely to be epenthesized by a general phonetic process to break up a consonantal cluster.⁹ If they were lexically specified one by one, then it would be difficult to explain their consistency in terms of both segmental quality and prosodic condition.

Another possibility that should be rejected is that the segments of the base are first copied completely and then a marked vowel is replaced by an unmarked one to minimize segmental markedness violations. If we took this markedness approach, then we would fail to capture the generalization that holds across reduplicative morphemes like the above and non-reduplicative ones like the following, where *s.luh* is a verb root and *n* is the infix for nominalization:

(4) a.	s.luh	[səluh]	‘to shoot’
b.	sn.luh	[sɛnluh]	‘shooting’

Notice that *ɛ* and *ə* consistently appear in the designated environments, *whether or not they are created by reduplication*, and that the copy-and-replacement approach can only accommodate reduplicated forms. The phonetic epenthesis approach, on the other hand, can handle both reduplicative and non-reduplicative morphemes *uniformly*.¹⁰

If this phonetic epenthesis hypothesis is true, then it turns out that some types of Mon-Khmer reduplication are highly unusual in terms of both the base-reduplicant

(B-R) association pattern and the reduplicant's prosodic form. Usually, the reduplicant corresponds to a contiguous (sub)string of the base, and takes the form of an "authentic unit of prosody", such as a syllable with a vowel nucleus or a disyllabic foot, which is illustrated by the reduplicated forms in (5):

- (5) a. CV Reduplication in Nootka:
 č̣i-č̣ims-‘i:ŋ ‘hunting bear’ (Stonham 1994)
- b. CVC Reduplication in Lushootseed:
 cas-cast ‘branches’ (Hess 1966)
- c. CVCV Reduplication in Boumaa Fijian:
 buta-butaʔo ‘steal on a number of occasions’ (Dixon 1988)

In contrast, some types of Mon-Khmer reduplication produce an apparently non-authentic unit of prosody, and it does not correspond to a contiguous substring of the base. For example, consider the first (prefixal) type of reduplication in (3a). The reduplicant of this type consists only of copies of the first and last consonants of the base; e.g., the first consonant *k* of the reduplicant *kh* in the word *kh.koh* corresponds to the first consonant of the base *koh*, and the last consonant *h* of the reduplicant to the last consonant of the base. Moreover, this reduplicant appears prosodically defective in that it has no vowel or sonorant consonant on which a basic syllable could be erected. The second (infixal) and fourth (stem-prefixal) types of reduplication also appear unusual in terms of the B-R correspondence and the resulting prosodic structure, although the third (prefixal with *a*) and the fifth (arguably suffixal) are familiar types of reduplication that are attested in many other languages.

The present study thus addresses the following questions, focusing primarily on the apparently unusual types of reduplication in Mon-Khmer languages:

- (6) a. Why is it that Mon-Khmer reduplication takes the form that it does?
 (In the case of Type 1, for example, why is it prefixal, and why not, say, suffixal?
 Why is it that only the first and last consonants are reduplicated, and why not
 other segments?)
- b. Why is it that Mon-Khmer languages have such apparently unusual types of
 reduplication, whereas many other languages (say, Japanese) do not?

In answer to these questions, this paper makes the following four claims. First, the apparently unusual types of Mon-Khmer reduplication, in fact, fall within the range of the regular prosody of the language, and even reflect its canonical prosodic structure. It turns out, therefore, that Mon-Khmer reduplication can be described as an instance of the emergence of the unmarked in the particular sense stated above: the independently-motivated, canonical prosodic structure of a language is recycled for a

new function and reveals itself as a reduplicated form.¹¹

Second, the unmarked prosodic structure of Mon-Khmer languages is the so-called "sesquisyllabic" structure with final prominence, which is a disyllabic iamb with a "minor syllable" (a syllable with no vocalic nucleus): $[(C(C)-C\acute{V}C)_{F\downarrow}]_{PrWd}$. I take it as the canonical iamb on grounds to be presented later, and suggest that this unmarked prosodic structure is motivated by a general constraint on rhythmic grouping:

(7) Iambic/Trochaic Law:

- a. Elements contrasting in intensity naturally form groupings with initial prominence.
 - b. Elements contrasting in duration naturally form groupings with final prominence.
- (Hayes 1995: 80)

This is a "domain-general" and "system-independent" constraint - "domain-general" in the sense that it pertains to human perceptual preference and applies to non-linguistic rhythmic grouping (e.g. music) as well as to phonological structure; "system-independent" in the sense that it gives preference to one structure over another, *independently of the state of a system to which they belong*. Under this system-independent characterization of (un)markedness, for example, the canonical iamb ($L\acute{H}$) is regarded as being less marked than an iamb with even duration like ($L\acute{L}$) *in any language*.

Third, the emergence of the unmarked in reduplication is observed not only in Mon-Khmer languages but also in many other languages, which will be illustrated later. The emergence of the marked, on the other hand, is exceedingly rare. Linguistic theory should accordingly be formulated in such a way that this robust crosslinguistic tendency follows from the interaction of general constraints, rather than that it is just described as an accident. It seems, however, that Optimality Theory, as it now stands, is not formulated this way. It basically permits any ranking of constraints, yielding a factorial typology.¹² It can thus account for absolute universals in terms of the (un)availability of particular constraint rankings, but not for universal tendencies. To make up for this limitation, this paper suggests that OT should be complemented with a set of "meta-constraints", higher-order constraints on the way of ranking constraints. *They characterize the (un)markedness of constraint ranking, in terms of which the (un)markedness of linguistic structure is characterized*. To accommodate the particular crosslinguistic tendency indicated above, the following meta-constraint is required:

- (8) Reduplication-specific constraints may not be ranked higher than constraints that are directly responsible for the canonical prosodic structure(s) of the language.

I assume that higher-order constraints like this, just like ordinary OT constraints, can

be violated, but only minimally and under duress. That is, they are implicitly qualified by "unless otherwise motivated". The meta-constraint in (8) therefore works practically as a "preferential constraint ranking", i.e., as a bias toward preferentially ranking certain types of constraints in a certain order. This paper shows that its interaction with a certain other general constraint gives rise to the crosslinguistic tendency in question, explaining why it is usually the case that what reduplicated form a given language has depends strongly on what prosodic structure it has. The above meta-constraint thus contributes to properly delimiting the range of possible/likely types of reduplication and in turn of grammars, thereby making the theory more restrictive.

It is, however, too restrictive in that it can only handle the particular linguistic dependency: the one between a prosodic structure and a reduplicated form. Other dependencies among linguistic entities, including the segmental dependency illustrated above, are beyond its scope. This paper thus goes a step further, suggesting a constraint on linguistic dependencies in general, from which (the practical effect of) the meta-constraint in (8) might be derived as a theorem. Specifically, it proposes a constraint requiring that linguistic entities share information with one another maximally, thereby explaining why they strongly tend to constitute a system with interdependent elements, rather than to remain independent of one another, and in turn why we strongly tend to make the best use of existing (sub)structures for new functions. The utilization of an independently-motivated prosodic structure for the purpose of reduplication is explained as an instance of it. This constraint, in contrast to the Iambic/Trochaic Law, is "system-dependent" in nature; for what entity is regarded as being unmarked depends crucially on the state of a system to which it belongs, particularly on what other entities it has. Under this system-dependent characterization of (un)markedness, what structure is unmarked may vary from language to language, and what structure is likely utilized for a new function as the unmarked may accordingly differ depending on the language.

In short, this paper argues that the apparently unusual properties of Mon-Khmer reduplication are due to four factors: (i) the canonical prosodic structure of the language, which is in turn attributed crucially to (ii) a domain-general and system-independent constraint on rhythmic grouping, and (iii) a meta-constraint on the way of ranking constraints on prosody and reduplication, which might be derived as a theorem from (iv) a system-dependent constraint on linguistic dependencies in general.

The arguments proceed as follows. Section 2 examines previous analyses of Mon-Khmer reduplication (both traditional and OT), pointing out that they are insightful but have some limitations in terms of both the breadth of descriptive

coverage and deductive depth. Section 3 discusses the problem with OT touched upon above, illustrating it with various types of reduplication in the world's languages that manifest the emergence of the (un)marked. Section 4 describes the regular prosodic structures of Mon-Khmer languages, spelling out their characteristic properties. Based on the descriptive generalizations, section 5 develops an alternative analysis of Mon-Khmer prosody and reduplication, giving answers to the main questions posed above. Finally, section 6 discusses the emergence of the unmarked and linguistic dependency in a broader perspective, exploring their origins as a direction for future research.

2. Previous Analyses

There are several previous studies that deal with Mon-Khmer reduplication; some are purely descriptive works (Benjamin 1976; Diffloth 1976a, b; Svantesson 1983); others give an account of it in terms of a particular phonological theory (McCarthy 1982; Broselow and McCarthy 1983; Sloan 1988; Shaw 1993, 1996; Takeda 1997; Gafos 1998; Hendricks 2001). The present study owes much to their descriptions and insights. This section reviews two major previous analyses of reduplication in Temiar and Semai,¹³ thereby setting out the scene for the main discussions to be presented in the subsequent sections: (i) Sloan's (1988) traditional (template and association) analysis, and (ii) Gafos's (1998) non-templatic OT analysis.

2.1 A Traditional Analysis: Sloan (1988)

Sloan (1988) deals with three types of Mon-Khmer reduplication: (i) C_1C_f prefixation to the heavy syllable of the base, (ii) C_f infixation immediately to the left of the heavy syllable of the base, and (iii) C_1C_f prefixation to the stem base. They correspond respectively to Types 1, 2 and 4 classified in (3). She makes no distinction between them in terms of reduplication process, and argues that the difference between the two types of prefixal reduplication (Types 1 and 4) results from the size of the base, whereas the difference between prefixal and infixal reduplication of the same base size (Types 2 and 4) results from the underlying form.

Let us consider how each type of reduplication is derived in Sloan's system, beginning with the prefixal reduplication. She assumes that the prefixal reduplication (of Type 4) in Semai (e.g. *gp.g.hup*) is generated by the following four-step derivation:

(9) Derivation of Semaï prefixal reduplication:

- a. Affixation
- s1 + s0 s
- | / | \
- g h u p
- i. s1 ii. s0
- C C C
- b. Copy
- s1 + s0 s
- | / | \
- ghup g h u p
- c. Special Association Principle (SAP): As the first step of association, associate the rightmost element of the copy to the affixal template.
- s1 + s0 s
- | / | \
- ghup g h u p
- d. Association (L → R)
- s1 + s0 s
- | / | \
- ghup g h u p
- s1 + s0 s
- | / | \
- ghup g h u p

The first step is affixation of what she calls s1 before s0, where s1 represents a syllable with two consonants only (onset and coda), and s0 a syllable with only one consonant (onset), as shown in (9ai) and (9aii), respectively. The second step is to copy the melody of the base (e.g. *g.hup*). The next step is crucial: the SAP in (9c) applies, thereby associating the last consonant of the copy to the s1 template. Finally, the normal association rule, which is left-to-right in the case of a prefix, applies and associates the leftmost consonant of the copy to the template. She argues that the derivation of the other type of prefixal reduplication (Type 1) proceeds in exactly the same way, assuming that the difference between the resulting forms (**CC.C.CVC** vs. **CC.CVC**) simply reflects the difference between the original bases (**C.CVC** vs. **CVC**).

On the other hand, infixal reduplication (Type 2) in Temiar (e.g. *sg.lɔg*) is derived in the following way:

(10) Derivation of Temiar infixal reduplication:

- a. Affixation
- s1 + s0 s
- | / | \
- s l ɔ g
- i. s1 ii. s0
- C C C
- b. Copy
- s1 + s0 s
- | / | \
- s l ɔ g
- c. SAP
- s l ɔ g s
- | / | \
- s l ɔ g
- s1 + s0 s
- | / | \
- s l ɔ g
- d. Association (L → R)
- s1 + s0 s
- | / | \
- s l ɔ g
- s1 + s0 s
- | / | \
- s l ɔ g

The crucial difference between this type and the previous one is that the initial consonant of the base is underlyingly syllabified in the previous case, but not in this case. It is therefore syllabified as the onset of *s1* in (10), while it is not in (9), where the onset of *s1* is not the initial consonant of the base but is that of its copy.

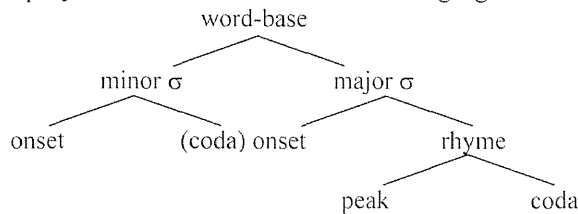
To be sure, these derivations yield the right reduplicant in the right position as far as these three types of reduplication are concerned. But here arise three questions, two of which correspond to the general questions posed earlier: (i) Can this analysis correctly extend to other attested types of reduplication in Mon-Khmer languages? (ii) Can it explain why Mon-Khmer reduplication takes the form that it does? (cf. (6a)) (iii) Can it explain why Mon-Khmer languages have the apparently unusual types of reduplication whereas many other languages do not? (cf. (6b)) The first question is descriptive, and the other two are explanatory. The answers to these questions are all in the negative, at least as her analysis now stands.

Let us consider them in turn, beginning with the first question. Notice first that Sloan takes an example of prefixal reduplication from Semai and an example of infixal reduplication from Temiar.¹⁴ In fact, however, Temiar and Semai have *both* prefixal and infixal reduplication, as shown in (3). Thus, the question to be addressed is whether or not her analysis can accommodate these two types of reduplication while distinguishing them from each other *within one language*. The answer seems to be no. Sloan attributes the differences between them to the differences in underlying representation between their bases. However, how can we reasonably decide what underlying representation a given lexical item has? If it were decided item by item (e.g., the initial consonant of *g.hup* is underlyingly syllabified and that of *s.lɔg* is not), then the generalization that holds among reduplicated words of each type would not be captured: reduplicated words of Type 4 are *all* underlyingly syllabified, and ones of Type 2 are *all* not. Moreover, the argument based on underlying representation is circular, because such a distinction between lexical items in terms of underlying representation has no other motivation than for classifying the types of reduplication in the first place.

Another descriptive limitation of Sloan's analysis is that it does not take into account the semantic difference between the two types of prefixal reduplication (Types 1 and 4). It thus fails to explain why reduplicated words derived by allegedly the same morphological operation depict rather different meanings: some (e.g. *kh.koh*) depict continuative aspect, and others (e.g. *gp.g.hup*) "expressive".¹⁵ Still another limitation is that it does not pay attention to reduplication of Types 3 and 5 (e.g. (3c) and (3e)). To accommodate those types of reduplication, a different affixation and/or association rule would be needed.

Let us now turn to the explanatory questions. Answering those *why*-questions, I argue, requires capturing the dependency relation between the basic prosodic structures and reduplicated forms of Mon-Khmer languages. It appears that she recognizes this prosodic dependency and attempts to capture it. The particular analysis she proposes, however, is not formulated in a way that only yields the apparently intended result. More specifically, by assuming a language-particular template and rule, she tries to conform Mon-Khmer reduplication forms to the characteristic prosodic structures of the language, which are called "sesquisyllabic structures". They are composed of "minor" and "major" syllables, and the "minor" syllable consists only of one or two consonants:

(11) Sesquisyllabic structures of Mon-Khmer languages:¹⁶



This insight behind the templatic analysis is well-taken, and will be incorporated into the alternative analysis to be presented in section 5. A limitation of her analysis, however, is that it has no general constraint that *ensures* that reduplicated forms fall within the range of the regular prosody of the language. This limitation might be attributed not to her particular analysis but rather to the theory on which it is based. In any case, many types of reduplication including unattested ones could be incorporated into a grammar without such a constraint. For example, recall that she stipulates a special syllabic template (s1 in (9) and (10)), its prefixation before the base, and a special association rule. But under her analysis or the theory on which it is based, one could instead assume, say, that the minor syllable template is suffixed to the major syllable. No such type of reduplication, however, is permitted in Mon-Khmer languages.

Given this claim, one might argue that this problem can be solved by assuming a *word* template like (11), instead of or in addition to a syllable template like the one in (9/10). To be sure, the traditional theory allows a language to have a word template, but, as will be pointed out shortly, it does not have a general constraint that requires a given language to formulate the template in terms of legitimate prosodic units in *that* language, let alone a constraint that motivates a given language to select a particular template preferentially out of many options. Thus, even an illegitimate structure in a language (say, CVC-CC in

a Mon-Khmer language) could be posited as a template without violating any constraint of the theory (as long as it is an "authentic unit of prosody", which will be discussed shortly). In short, the problem is that her analysis or the theory on which it is based is not formulated in such a way as to exclude such unattested types of reduplication by a general constraint, although it could rule them out by stipulating a language-particular template and/or rule. It thus seems insufficient to give a reasonable answer to the first explanatory question: Why is it that Mon-Khmer reduplication takes the form that it does?

Moreover, her analysis/theory seems to have no constraint that prohibits the minor syllable template from being used in languages without such a syllable. It thus fails to exclude the possibility that even languages without such an apparently unusual syllable (say, Japanese) may have reduplicated forms of that form. No languages have been attested, however, which have a minor (CC) syllable reduplicant without such a syllable in the regular prosody. In general, an "authentic unit of prosody" in one language is not necessarily so in another, but this observation may not be captured if the crucial notion is only characterized in universal terms. That is, it is not sufficient to simply list the members of "authentic units of prosody" in Universal Grammar, and to formulate the Prosodic Morphology Hypothesis in terms of them, as follows:

(12) Authentic units of prosody: {Prosodic Word, Foot, Syllable, Mora}

(13) Prosodic Morphology Hypothesis:
Templates are defined in terms of the authentic units of prosody.

It thus seems that the second explanatory question as well as the first one is yet to be answered: Why is it that Mon-Khmer languages have the apparently unusual types of reduplication whereas many other languages do not?

There are (at least) two ways of solving these problems. One is to assume general constraints on the way of formulating templates. The other is to abandon the use of language-particular templates altogether and instead explain their effects in terms of universal constraints. The current studies on phonology and prosodic morphology are pursuing the second approach, to which let us turn (cf. McCarthy and Prince 1995, 1999; Urbanczyk 1996, 1998; Spaelti 1997; Gafos 1998; Walker 2000; Hendricks 2001; Kennedy 2002; Crowhurst 2004).¹⁷

2.2 An OT Analysis: Gafos (1998)

Gafos (1998), as well as Sloan (1988), recognizes the dependency relation between

the sesquisyllabic structures and reduplicated forms of Mon-Khmer languages (particularly of Temiar), and attempts to capture it.¹⁸ His attempt succeeds to a larger extent than Sloan's does, since he formulates several prosodic constraints generally enough, attributing the prosodic form of Temiar reduplication to the independently motivated, regular prosody of the language. He explicitly states that the structure of Temiar reduplication "follows from the interaction of independent constraints on the prosody of the language" (p. 238). I agree with him on this point, and reach the same conclusion. Overall, his analysis is rigorous and insightful, and the present study is based largely on his as well as Sloan's descriptions and insights.

It should be pointed out, however, that his analysis also has some limitations. In what follows, I first show how and to what extent his analysis can account for Mon-Khmer reduplication, and then point out that it shares some limitations with Sloan's.

Gafos (1998) focuses on reduplicated forms of Temiar verbs, which are inflected as follows:

(14) Verb inflections in Temiar:

	Biconsonantal root	Triconsonantal root
Active voice:		
a. Perfective:	kɔ:w	s.lɔg
b. Simulfactive:	ka.kɔ:w	sa.lɔg
c. Continuative:	kw.kɔ:w	sg.lɔg
Causative voice:		
d. Perfective:	tr.kɔ:w	sr.lɔg
e. Simulfactive:	t.ra.kɔ:w	s.ra.lɔg
f. Continuative:	t.rw.kɔ:w	s.rg.lɔg

The reduplication exemplified in this paradigm can be categorized into the following three types, which correspond to the first three types classified in (3): (i) C_1C_f prefixation (e.g. *kw.kɔ:w*), (ii) C_1 prefixation with the fixed vowel *a* (e.g. *ka.kɔ:w*), and (iii) C_f infixation (e.g. *sg.lɔg*).

Gafos's analysis is based on McCarthy and Prince's (1995) correspondence theory, and assumes the following constraints on the identity between the base and the reduplicant:

(15) B-R identity constraints:

- a. Max-BR: Every segment of B has a correspondent in R.
- b. Dep-BR: Every segment of R has a correspondent in B. (p. 226)
- c. SRole: A segment in R and its correspondent in B must have identical syllable roles. (p. 227)
- d. Align (Affix, R, σ, L): The right edge of Affix must be aligned with the left edge of the stressed syllable, where 'Affix' ranges over the set {Simulfactive, Continuative}. (p. 237)

In addition to these B-R identity constraints, he assumes several phonological and input-output (I-O) faithfulness constraints, of which the following are particularly relevant to the present discussion:

(16) Phonological constraints:

- a. Onset: Every syllable has an onset. (p. 236)
- b. *Prefinal-V: Prefinal (= unstressed) vowels are not allowed. (p. 235)
- c. Markedness(*PL/ χ): A segment whose place of articulation is more marked than χ may not be used. (p. 241)

(17) I-O faithfulness constraint:

Max^{AFFIX-IO}: Every segment of the affix in the input must have a corresponding segment in the output. (p. 239)

If these constraints are ranked in the following way, he argues, the optimal candidates win out:

- (18) Max^{AFFIX-IO}, Onset, Align (Affix, R, $\acute{\sigma}$, L), SRole, Dep-BR >> *Prefinal-V >> Markedness >> Max-BR

Let us see in turn how this system works for the three types of Temiar reduplication, beginning with the first type: $C_1C_F.C_1V(:)C_F$ (e.g. *kw.kɔ:w*). The crucial constraints for this type of reduplication are *Prefinal-V, Align (Affix, R, $\acute{\sigma}$, L), and Max-BR. Assuming that the first two constraints outrank the last one, then a vowelless reduplicant correctly emerges as a prefix at the expense of the lower-ranking B-R identity constraint:

(19)

[Input: α^{RED} , c^1vc^2]	*Prefinal-V	Align	Max-BR
a. $c^1v.c^1vc^2$	*!		*
b. $c^1vc^2.c^1vc^2$	*!		*
c. $c^1vc^2.c^1c^2$		*!	*
d. $\text{☞} c^1c^2.c^1vc^2$			*

The first two candidates have a vowel in the reduplicant. They violate *Prefinal-V, because stress always falls on the final syllable in Mon-Khmer languages (cf. Benjamin 1976, Thomas 1992, Miller and Miller 1995). The third candidate with a suffixal reduplicant breaches the alignment constraint. On the other hand, the fourth candidate respects both the phonological and alignment constraints, although it violates the B-R identity constraint

Max-BR, due to the failure to copy the vowel of the base. Since Max-BR is ranked lower than the other two constraints, the last candidate emerges as the optimal form.

Let us next consider the prefixal reduplication with the fixed vowel *a*: $C_1a.C_1V(:)C_f$ (e.g. *ka.kɔ:w*). Gafos argues that this type follows from the constraint ranking: $\text{Max}^{\text{AFFIX-IO}}$, $\text{Ons} \gg \text{Markedness} (*\text{PL}/c^2) \gg \text{Max-BR}$.

(20)

[Input: a^{RED}, c^1vc^2]	$\text{Max}^{\text{AFFIX-IO}}$	Ons	Markedness	Max-BR
a. $c^1.c^1vc^2$	*!		$*(c^1)$	**
b. $c^1_c^2.c^1vc^2$	*!		$*(c^1) *(c^2)$	*
c. $a.c^1vc^2$		*!		***
d. $\text{☞} c^1a.c^1vc^2$			$*(c^1)$	**
e. $c^1ac^2.c^1vc^2$			$*(c^1) *(c^2)!$	*

The first two candidates violate $\text{Max}^{\text{AFFIX-IO}}$ since the prespecified input vowel *a* is not present in the output. The candidate with an onsetless syllable in (c) runs afoul of the other top-ranking constraint *Ons*. Thus, all of these three candidates lose to the other two, which obey the undominated constraints. Candidate (d) incurs one less violation of *Markedness* than candidate (e) does, since it copies one less segment than (e) does, at the expense of the lower-ranking constraint *Max-BR*. It thus turns out to be the winner.

What should be noted here is that the decisive constraint in selecting the CV reduplicant in (d) over the CVC reduplicant in (e) is *Markedness*, and that it is a constraint on segmental contents, and not on prosodic structures. What seems to matter, however, is not which consonant is copied in the reduplicant, but is rather what prosodic shape the reduplicant (and in turn the reduplicated word as a whole) takes. If what matters were the segmental content of a copied consonant, then it would be predicted that there exist cases where, for example, two CVC reduplicants behave differently depending on the segmental property of the coda consonant. However, no such cases have been attested in Temiar. The analysis based on segmental markedness, therefore, seems unrevealing.¹⁹

Let us turn to the infixal reduplication: $C_1C_f.CV(:)C_f$ (e.g. *sg.lɔg*). Gafos argues that the position of the copied element is determined by *Align* (*Affix*, *R*, *σ*, *L*), and its segmental property by constraints on the regular prosody of the language: **Prefinal-V* and *Markedness*.²⁰

(21)

[Input: α^{RED} , s.lɔg]	*Prefinal-V	Align	Markedness	Max-BR
a. α^{sg} .lɔg			*(g)	***
b. sɔ.lɔg	*!			***
c. gs.lɔg		*!	*(g)	***
d. sg.s.lɔg		*!	*(s) *(g)	**
e. s.lɔg.lɔg	*!		*(l) *(g)	*

The first candidate should be optimal for this type of reduplication, and it seems to win out correctly in the system. Candidate (b) has a vowel in the reduplicant, thus incurring a fatal violation of *Prefinal-V. Candidate (c) violates Align since the consonantal affix is not aligned with the left edge of the major syllable, on which stress falls. Candidate (d) is ruled out for the same reason. Assuming that the base in candidate (e) gets stressed, then the infixed reduplicant with a vowel would incur a violation of *Prefinal-V and lose against candidate (a).²¹ It thus follows that this analysis predicts that Temiar would not have reduplicated words of the types in (d) and (e): $C_1C_r.C_1.CVC_r$ and $C_1.CVC_r.CVC_r$. In fact, however, Temiar *does* have these types of reduplication, as illustrated in (3d) and (3e), respectively (although the CVC reduplicant is regarded as a suffix rather than an infix).

In short, Gafos's analysis can handle three types of Temiar reduplication (Types 1-3 in (3)), but not the other two (Types 4 and 5): it predicts that Temiar would not have them.²² The use of the segmental constraint Markedness in the accounts of Types 2 and 3 also needs to be reconsidered.

Over and above these empirical/descriptive limitations, theoretical/explanatory problems should also be noted about Gafos's OT analysis. Firstly, even when only the correct candidate is derived from the interaction of certain constraints ranked in a certain order, the particular ranking of the particular constraints does not follow from any general constraint of OT. That is, OT does not explain why they are ranked in the way they are. It thus allows many other rankings as well as the actually attested ones. It even allows reduplication-specific constraints to outrank constraints that are directly responsible for the canonical prosodic structures of the language. This ranking in turn yields the emergence of a structure in reduplication that is not permitted in the regular prosody of the language. As it turns out in the next section, however, this is inconsistent with the extensive observations of reduplication in the world's languages.

Let me here illustrate this potential problem by a hypothetical example of reduplication. Recall first that Gafos takes the base of the reduplication in Temiar as a

stressed syllable (recall his formulation of the alignment constraint in (15d)). But OT (or the Generalized Alignment format) needs to allow many other types of alignment constraints to accommodate reduplication in other languages (McCarthy and Prince 1993b). For instance, as pointed out above, Sloan (1988) claims that the CC reduplicant is prefixed to a morphological stem in Semai (e.g. *gp.g.hup*), and Temiar also has this type of reduplication.²³ We thus need to posit an alignment constraint of the following type, and to rank it high enough to exhibit its effect in Mon-Khmer languages:

- (22) Align (Affix, R, Stem, L): The right edge of Affix must be aligned with the left edge of the stem, where Affix is expressive (prolongation or continuous repetition in time).

Moreover, given many languages with total reduplication, Max-BR also needs to be assumed and ranked high enough. Here arises a potential problem with OT: unless it has a meta-constraint on the way of ranking constraints, it would allow the possibility that the two constraints (Align_{Affix} and Max-BR) are both ranked high, say, higher than *Prefinal-V in a language where no major syllable can appear before a minor syllable within a prosodic word. In this hypothetical language (let us call it Temiar'), the reduplicant would be a full copy of a morphological stem and appear before it. Given a stem of the form (C.CVC), for example, the reduplicated form would be C.CVC.C.CVC, despite such a structure being prohibited in the regular prosody of the language:

- (23) Reduplication in Temiar' (a hypothetical language):
 $C_1.C_2V(:)C_f \rightarrow C_1.C_2V(:)C_f.C_1.C_2V(:)C_f$
 a. *s.log* → **s.log**.s.log (a'. ***s.log**.s.log in Temiar)
 b. *k.lã:d* → **k.lã:d**.k.lã:d (b'. ***k.lã:d**.k.lã:d in Temiar)

This type of reduplication thus represents the emergence of *the marked or more marked than the language as a whole*. Reduplication of this kind, in fact, is not acceptable in the actual language Temiar. Does this just happen to be the case in this particular language, or does it reflect a robust crosslinguistic tendency to avoid marked structures and to seek unmarked structures as long as communicative functions are met? If the former turns out to be the case, then the problem would not be serious, and no theoretical revision might be required. The solution would be to just stipulate a particular ranking of the relevant constraints, as is customarily done in OT. If the latter turns out to be the case, however, linguistic theory should capture it, and a theory that fails to do it should accordingly be modified. Let us see which is in fact the case in the next section, and

then return to this problem and reconsider its theoretical significance.

3. The Emergence of the (Un)Marked

This section shows that the emergence of the unmarked or less marked than the language as a whole is pervasively observed, whereas that of the marked or more marked is exceedingly rare. It also points out that the previous OT analyses of reduplication, including Gafos (1998), fail to capture this proportion, and should be revised accordingly.

Let us first consider some examples of partial reduplication that represent the emergence of an unmarked or less marked syllable structure than the language as a whole:^{24, 25}

(24) Emergence of CV:

- | | | |
|----------------|--|-----------------------------------|
| a. From CVC: | | |
| i. | met- go -got ²⁶ ‘strong’ | [Chamorro] (Topping 1980) |
| ii. | si-sil’ ‘calico (diminutive)’ | [Thompson] (Haerberlin 1918) |
| iii. | so-songo ‘tree’ | [Kela] (Steriade 1988) |
| iv. | may- ma ‘separate, apart’ | [Kashaya] (Buckley 1994) |
| v. | wi-winni ‘to stand’ | [Southern Paiute] (McCarthy 2002) |
| vi. | ʔa-ʔalʔal ‘hut’ | [Lushootseed] (Urbanczyk 1996) |
| vii. | si- pa -pandilij ‘wearing a skirt’ | [Ilokano] (Hayes and Abad 1989) |
| b. From CVV: | | |
| i. | ta- taa ‘strike’ | [Samoan] (Marantz 1982) |
| ii. | ʔa-ʔaaral ‘to read’ | [Tagalog] (Cabrera et al. 1965) |
| iii. | too- to ‘dance movement’ | [Kashaya] (Buckley 1994) |
| iv. | juu- ju -k ‘tender’ | [Winnebago] (Miner 1979) |
| v. | ku-kui ‘mesquite tree’ | [Papago] (Fitzgerald 2000) |
| vi. | ta- toa ‘treading/kicking’ | [Nakanai] (Johnston 1980) |
| vii. | kû-kûelò ‘nibble at’ | [Mbe] (Walker 2000) |
| c. From CVVC: | | |
| i. | la- laud ‘big, old’ | [Ponapean] (Rehg and Sohl 1981) |
| ii. | li-lòir ‘clean-3s’ | [West Tarangan] (Nivens 1993) |
| iii. | pi- piid ‘seeing’ | [Papago] (Fitzgerald 2000) |
| d. From CVCC: | | |
| i. | ma- mand ‘tame’ | [Ponapean] (Rehg and Sohl 1981) |
| ii. | go- gogs ‘dog’ | [Papago] (Fitzgerald 2000) |
| e. From CCVC: | | |
| | sa-svar-a ‘sound’ | [Sanskrit] (Steriade 1988) |
| f. From CCVCC: | | |
| | sa-syand-a ‘move on’ | [Sanskrit] (Steriade 1988) |

(25) Emergence of CVV:

- | | | |
|--------------|--------------------------------------|----------------------------|
| a. From CVVC | | |
| i. | waa -waas-čit’ ‘naming where’ | [Nootka] (Stonham 1994) |
| ii. | čaa -čaak ‘bend’ | [Mokilese] (Harrison 1976) |
| b. From CCVC | | |
| | saa-svap ‘sleep’ | [Sanskrit] (Steriade 1988) |

(26) Emergence of CVC:

a. From CVCC:

- i. **cas**-cast ‘branches’ [Lushootseed] (Hess 1966)
- ii. **mar**-mard ‘rub, crush’ [Sanskrit] (Steriade 1988)
- iii. **tərg-tər** ‘meat’ [Palan Koryak] (Spencer 1991)

b. From CCVC:

- i. **tan**-stan ‘thunder’ [Sanskrit] (Steriade 1988)
- ii. **stan-tan**-mc ‘all one’s female ancestors on the mother’s side’
[Bella Coola] (Nater 1984)

c. From CVVC:

- qap**-qaap ‘to scratch something’ [Nisgha] (Tarpent 1983)

d. From CCVCC:

- kan**-i-skand ‘leap’ [Sanskrit] (Steriade 1988)

(27) Emergence of CC_{son}:a. From CC_{son}V:

- sm**-sma ‘to tell a story, myth’ [Bella Coola] (Nater 1984)

b. From CC_{son}C:

- tl**-tlk^w ‘swallow (continuative)’ [Bella Coola] (Newman 1971)

c. From CC_{son}VC:

- q^wl**-q^wlay ‘logs’ [Lushootseed] (Broslow 1983)

d. From CC_{son}VVC:

- x^wn**-x^wnaal-i ‘spring of water (diminutive)’ [Bella Coola] (Newman 1971)

These various types of partial reduplication all exemplify the emergence of an unmarked or less marked syllable structure as the reduplicant. In (24), for example, the most basic syllable structure CV emerges as the reduplicant from the base CVC (24a), CVV (24b), CVVC (24c), CVCC (24d), CCVC (24e), or CCVCC (24f). In (25), CVV emerges from CVVC (25a) or CCVC (25b). The emergent syllable CVV might be more marked than CV in terms of the complexity/length of the nucleus, but is less marked than the closed syllable base in that it has no coda. The reduplicant in (26) takes the form of CVC, but is less marked than the base with a complex coda (26a), with a complex onset (26b), with a complex nucleus (26c), or with two complex margins (26d). The reduplicant syllables in (27) are all erected on a sonorant consonant, but note that CC_{son} is a legitimate syllable structure in the Salish languages (e.g., Bella Coola and Lushootseed). Note also that it has a simple onset and no coda, thus being less marked than the syllables with a complex onset/nucleus and/or a coda in (27a-d). Therefore, even the reduplicated words in (27) can also be taken as instances of the emergence of a less marked syllable structure than the language as a whole.

The emergence of a metrically unmarked or less marked structure as the reduplicant has also been attested in many languages:

(28) Emergence of the minimal word: $[(\sigma\sigma)_{Fi}]_{PrWd}$ or $[(\mu\mu)_{Fi}]_{PrWd}$

a. From a base longer than it:

- | | | |
|---|-----------------|-----------------------|
| i. kintal -kintalpa ‘lizard species’ | [Yidij] | (Dixon 1977) |
| ii. ʔilpa -ʔilparku ‘bird species’ | [Diyari] | (Austin 1981) |
| iii. hite -hiteali ‘laugh’ | [Paamese] | (Russell 1997) |
| iv. salaga-laga ‘long’ | [Manam] | (Lichtenberk 1983) |
| v. koo -kooko ‘grind coconut’ | [Mokilese] | (Harrison 1976) |
| vi. kal -kaldij ‘goats’ | [Ilokano] | (Hayes and Abad 1989) |
| vii. kar -karkara ‘to snore’ | [Rotuman] | (Churchward 1978) |
| viii. maa -maarau ‘be permanently happy’ | [Boumaa Fijian] | (Dixon 1988) |

b. From a base shorter than it:

paa -pa ‘weave’	[Ponapean]	(Rehg and Sohl 1981)
------------------------	------------	----------------------

The examples in (28a) show that if the base is longer than a disyllabic or bimoraic foot, then it is shortened by less-than-full copying into the metrically unmarked structure: $[(\sigma\sigma)_{Fi}]_{PrWd}$ or $[(\mu\mu)_{Fi}]_{PrWd}$. On the other hand, if the base is too short to satisfy the minimal word condition, then the vowel is lengthened to make up for the failing, as shown in (28b). This Ponapean example represents a case alluded to earlier (note 25): syllabic unmarkedness is in conflict with metrical (foot or prosodic word-level) unmarkedness, and the conflict is resolved in such a way that the former gives way to the latter (cf. Kennedy 2002) ^{27, 28}

This ubiquity of the emergence of the unmarked suggests that it reflects an essential property of language, which should be captured by linguistic theory one way or another. Certainly, it has been recognized by the proponents of OT. They have not simply described it, but have also accounted for it elegantly, as the emergence of the unmarked in the technical sense: the effect of a constraint emerges when a higher-ranking constraint is not applicable.

Let us consider, more specifically, how the OT account of the emergence of the unmarked goes, using the reduplication of the type in (24a) as an example. The fact that the closed syllable is permitted in the base indicates that No-Coda may be violated in the language as a whole, and that it should be ranked lower than the I-O faithfulness constraints. Given this ranking, the input CVC base may surface as is, respecting the dominating faithfulness constraints while sacrificing the dominated prosodic constraint. Thus the effect of No-Coda is not always visible in the base. However, it could emerge in the reduplicant where the higher-ranking I-O faithfulness constraints are not at issue; for the reduplicant by definition has no underlying representation. It does happen iff a B-R identity constraint like Max-BR is ranked lower than the prosodic constraint. This ranking allows a candidate with the coda-sparing reduplicant CV to be judged better than one with the complete copying of the base CVC. As shown in (29), the partial reduplication form incurs one less violation of the higher-ranking prosodic constraint than the total reduplication form does:

(29) Emergence of the effect of No-Coda in reduplication

	I-O Faithful	No-Coda	Max-BR
CV-CVC		*	*
CVC-CVC		**!	

This is a paradigm case of the emergence of the unmarked in an OT grammar. Thus, OT can provide an elegant account of the emergence of the unmarked phenomenon with the following skeletal constraint ranking (McCarthy and Prince 1994, 1995, 1999):

- (30) I-O Faithfulness >> Phono-Constraint >> B-R Identity

It should be noted, however, that OT permits other ranking patterns including the following *as equally likely possibilities*:²⁹

- (31) B-R Identity >> Phono-Constraint >> I-O Faithfulness

This ranking predicts that even a structure which would be ruled out by a phonological constraint could emerge in reduplication if (and only if) a higher-ranking B-R identity constraint enforces it. That is, this ranking allows the emergence of *the marked*.

The question is: Is the emergence of the marked attested as widely as that of the unmarked is? The answer is no: it is exceedingly rare. As far as prosodic markedness is concerned, one exceptional case that reportedly represents the emergence of the marked is reduplication in Mangap-Mbula (an Austronesian language). Spaelti (1997) observes some examples where a CVC syllable emerges as a prefixal reduplicant, although closed syllables are otherwise not permitted in word-initial/internal positions in this language:

(32) Reduplication in Mangap-Mbula

- a. **bad**-'baada 'you (sg) be carrying'
- b. **mot**-'mooto 'worms'
- c. i-**tor**-'tooro '3sg-turn'

These examples thus indicate that the emergence of the marked is not prohibited altogether in human languages.

This does not mean, however, that any marked structures are freely allowed to emerge in reduplication. In the case at hand, for example, CVC syllables are not *completely* prohibited in Mangap-Mbula; they *do* occur in the word-final position, as shown below:

- (33) a. posop 'you sg. finish'
 b. timeⁿder 'they stand'
 c. tipo^mbol 'they cause-be strong'

Their emergence as reduplicated forms, therefore, should be taken as an instance of the emergence of a *slightly* (rather than a great deal) more marked structure than the language as a whole. For comparison, imagine a case where a very complex structure like, say, CCCVCC suddenly emerges as a reduplicated form in a language where no syllable other than CV is permitted in any position. Of course, no such case has been attested in the world's languages. Proper descriptive generalizations thus turn out to be the following: *the emergence of the unmarked is pervasively observed, whereas that of the marked is exceedingly rare; if any, only a slightly marked structure can make its appearance.*

In short, a limitation of OT is that it incorrectly predicts that the emergence of the marked is as likely to happen as that of the unmarked is, and fails to capture the observation that the overwhelming majority of partial reduplication is of the type that yields a less marked structure than the language as a whole. The theory should therefore be modified in such a way as to capture this robust crosslinguistic tendency. As indicated earlier, one way might be to assume a higher-order constraint on the way of ranking constraints. The simplest way of restricting constraint ranking possibilities might be to simply fix the emergence of the unmarked ranking in (30). It will be shown later, however, that this "solution" does not work as (30) now stands, and that it should be reformulated more specifically, in terms of canonical prosodic structures.

With this observation and consideration, going back to the specific problem left unsolved in the preceding section, we can see that it is a serious problem in that it is not so much a problem with the particular analysis (Gafos 1998) of the particular empirical phenomenon (Mon-Khmer reduplication), but is rather a problem with the general theory on which it is based: OT. The analysis based on such a theory allows illegitimate prosodic structures of a given language to emerge as reduplicated forms. This is because the theory is not restricted in a way that basically prohibits the B-R identity constraints (e.g. Max-BR) from outranking the constraints that are responsible for the canonical prosodic structures of the language (e.g. *Prefinal V). Thus, the hypothetical reduplication in (23) (e.g., **s.lɔg.s.lɔg*) has turned out to represent an unlikely type of phenomenon that would result from an unlikely constraint ranking pattern.

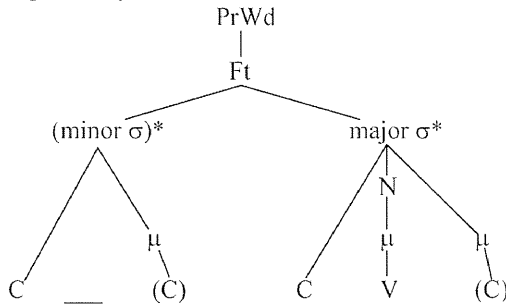
One more point to be noted about that hypothetical reduplication is that the reduplicant itself (e.g., *s.lɔg*) is prosodically legitimate in the actual language Temiar (it could stand by itself as a word), and what is illegitimate is rather the prosodic word as a whole including the base and the reduplicant (e.g., *s.lɔg.s.lɔg*). This should be contrasted with the examples of the emergence of the unmarked in (24) - (28), where the reduplicant itself (e.g., *go* in (24ai)) represents a prosodically unmarked structure. Conversely, in acceptable reduplicated

forms like CC-CVC (e.g., *kh.koh*), the bare-consonantal reduplicant CC appears prosodically illegitimate on its own, in the sense that it is not licensed by a higher unit in the prosodic hierarchy without attaching itself to a certain prosodic unit (we shall see in the next section that, in fact, it is).³⁰ Thus, a question to be addressed is: In what sense does Mon-Khmer reduplication represent the emergence of the unmarked if at all? I shall claim in section 5 that Mon-Khmer reduplication, particularly Type 1, is a rather rare type of the emergence of the unmarked phenomenon, in the sense that the overall structure that emerges as a reduplicated form (CC-CVC) is the canonical prosodic structure of the language, but it is created by adding to the base (CVC) a reduplicant that has an illegitimate prosodic structure on its own (CC). Making arguments for this claim, however, requires a close examination of the regular prosody of Mon-Khmer languages. This is the topic of the section to follow.

4. The Regular Prosody of Mon-Khmer Languages: Descriptive Generalizations

This section describes the regular prosodic structures of Temiar and Semai, specifying the range of their canonical and legitimate prosodic structures. I assume with Shaw (1993, 1996) that they have the following prosodic structures:³¹

(34) Legitimate prosodic structures of Temiar and Semai:



The "major syllable" is defined as one with a nuclear mora, and the "minor syllable" is one without it. Let us see in turn (i) the status and internal structure of the major syllable, (ii) those of the minor syllable, and (iii) their relationship.

The major syllable has at least five properties. First of all, it is obligatory in the sense that it is required for a word to stand: there are no words without a major syllable in Mon-Khmer languages. Second, an onset is obligatory: no major syllables without an onset are permitted. Third, a coda is basically optional, but is required in the word-final position in Temiar and Semai.³² Fourth, complex margins are prohibited in Temiar and Semai.³³ Finally, stress always falls on the word-final (heavy) syllable, a property that plays an important role in the

analysis of reduplication to be presented in the next section. Thus, the following structures are all legitimate in Temiar and Semai, independently of reduplication:

(35) CVC words:

- | | |
|-------------------|--|
| a. Temiar: | b. Semai: |
| i. jiʔ 'sick' | i. dər 'flame' |
| ii. tɛʔ 'earth' | ii. dic 'completed' |
| iii. doʔ 'to run' | iii. koh 'to chop off' |
| iv. tab 'egg' | iv. pɛc 'to throw, shoot' |
| v. tuh 'to tell' | v. dɔp 'to settle on a place off the ground' |

(36) CV:C words:

- | | |
|----------------------|-----------------------|
| a. Temiar: | b. Semai: |
| i. tu:k 'to fear' | i. ɲa:r 'to face' |
| ii. tɛ:ʔ 'just now' | ii. du:y 'evening' |
| iii. bə:h 'father' | iii. bi:t 'to squint' |
| iv. de:k 'house' | iv. ɡə:y 'to sit' |
| v. ɡə:b 'completely' | v. ɲə:c 'burnt' |

(37) CV.CV(:)C words:

- | | |
|------------------------------|---|
| a. Temiar: | b. Semai: |
| i. ha.lab 'to go down river' | i. ga.ra:c 'to slide' |
| ii. bo.ləh 'can' | ii. ja.lɛc 'small plant' |
| iii. ma.doh 'to here' | iii. pa.ɡɛc 'to sharpen a tool roughly/quickly' |
| iv. ʔi.ci:b 'I go' | iv. sa.ɡɔc 'to pry (e.g. a log) open half way' |
| v. do.rɛ:h 'downwards' | v. sa.lic 'to occupy, take over a place' |

(38) CVC.CV(:)C words:³⁴

- | | |
|------------------------------|--|
| a. Temiar: | b. Semai: |
| i. sin.dul 'to float' | i. kan.rɔɲ 'protruding muscles of thin person' |
| ii. num.naʔ 'from there' | ii. raɲ.kaɲ 'skeletal' |
| iii. diɲ.yɛ:w 'guard house' | iii. raɲ.ɡɔ:c 'bare-necked (as of chicken)' |
| iv. hum.bo:ʔ 'normally' | iv. run.tuc 'to snatch' |
| v. ʔun.tu:y 'they elsewhere' | |

The minor syllable has some characteristic properties, which make Mon-Khmer prosodic structures apparently unusual. First of all, it is "parasitic" in the sense that it cannot stand alone and requires a major syllable to constitute a word. There are no words that consist only of a minor syllable/syllables in Mon-Khmer languages. Second, as indicated earlier, the minor syllable is phonetically realized with a vowel, but its segmental quality is predictable from the prosodic environment around it. It is therefore likely that the vowel is not lexically specified but is rather epenthesized by a phonetic process. Consider the following pair of Temiar words for example:

- (39) a. s.lɔg [səlɔg] 'to lie down' (root/active voice, perfective aspect)
 b. sg.lɔg [sɛglɔg] 'to be lying down' (active voice, continuative aspect)

The perfective form *s.lɔg* is phonetically realized as *səlɔg*, but the continuative form as *sɛglɔg*. If the short mid vowels *ə* and *ɛ* were used only for these words, then one would assume that they are lexically specified. However, a few exceptions aside (recall note 9), they consistently appear in the same environment: *ə* in the open syllable, and *ɛ* in the closed syllable. Consider the following pairs of words, where *n* is the affix for nominalization:

- (40) a. sg.ləg [sɛgləg] 'to knot'
 b. s.ng.ləg [sənɛgləg] 'knot'
- (41) a. c.ɛr [cɛr] 'to pare'
 b. c.nɛr [cənɛr] 'knife'
- (42) a. s.luh [səluh] 'to shoot'
 b. sn.luh [sɛnluh] 'shooting'
- (43) a. go.lap [golap] 'to carry on shoulder'
 b. g.no.lap [gənolap] 'carrying on shoulder'
- (44) a. sin.dul [sindul] 'to float'
 b. s.nin.dul [sənindul] 'floating'

The Temiar verb *sgləg* 'to knot' in (40) is phonetically realized in the same way as the continuative form of the verb cited just above, which is similar to it prosodically and segmentally. If it is nominalized by the infixation of *n* in the second position and the resulting form is syllabified from right to left, then it is realized as *sənɛgləg*, as predicted by the above epenthesis hypothesis. This hypothesis can apply to other types of roots, whether they are biconsonantal (41), triconsonantal with a minor syllable (42), triconsonantal with no minor syllable (43), or quadriconsonantal (44).

The internal structure of the minor syllable is almost the same as that of the major syllable, except for the absence of a nuclear vowel. An onset is obligatory, but a coda optional. If it appears, then it is moraic. No complex margins are permitted. Thus, the structure of the minor syllable is always C_ or C_C, where the underlines indicate epenthesis positions. The epenthesized vowel is arguably not moraic, which will be discussed in the next section. Minor syllables, again like major syllables, can be concatenated with each other. Stress never falls on any minor syllables. This is presumably the reason for Thomas's (1992) observation that the syllabic status of the minor syllable is

sometimes unclear in poetry, chanting, and singing (thus words with a minor syllable are sometimes counted as one syllable, sometimes two); hence the name "sesquisyllabic".³⁵ Due to these properties of the minor syllable, the following sesquisyllabic structures with stress on the final syllable are all permitted in Temiar and Semai, *independently of reduplication*:

(45) C.CV(:)C words:

a. Temiar:

- i. s.ləg 'to lie down'
- ii. s.luh 'to shoot'
- iii. c.nɛr 'knife'
- iv. r.wa:y 'head-soul'
- v. k.ro:p 'underside'

b. Semai:

- i. m.mat 'shrub'
- ii. p.lɛ:ʔ 'fruit'
- iii. s.tit 'sling'
- iv. c.ŋa:l 'red'
- v. c.ʔɛ:t 'sweet'

(46) CC.CV(:)C words:

a. Temiar:

- i. sg.ləg 'to knot'
- ii. sn.luh 'shooting'
- iii. br.ca:ʔ 'to feed'
- iv. kr.wa:k 'to frame'
- v. gr.lut 'long and thin'

b. Semai:

- i. kr.lɛc 'to extract the pit of a fruit'
- ii. sn.ʔɔ:y 'human being'
- iii. mr.hũ:r 'snake'
- iv. gr.par 'mountain imperial pigeon'
- v. gr.pər 'small bat'

(47) C(C).CC/V.CV(:)C words:

a. Temiar:

- i. k.rn.wa:k 'frame'
- ii. t.rn.ʔɔj 'raising'
- iii. s.ŋg.ləg 'knot'
- iv. gn.gr.lut 'spindliness'
- v. g.no.lap 'carrying on shoulder'

b. Semai:

- i. b.rk.yɛ:k 'several white things'
- ii. k.lc.wɛc 'irregular flapping
circular movements'
- iii. r.rŋ.ʔaŋ 'appearance of irregular
cracks'
- iv. k.nm.ji:p 'feelings'
- v. k.la.tā:p 'the appearance of a
swollen ant-bite'

Finally, what should be noted about the relation between the major and minor syllables is that the minor syllable always precedes the major syllable within a prosodic word. Therefore, prosodic word structures like the following are all prohibited:

(48) Some illegitimate prosodic word structures of Temiar and Semai:

- a. *CVC.C.CVC
- b. *CVC.CC.CVC
- c. *CV.C.CC.CVC
- d. *C.CV.CC.CVC.C
- e. *CV.CC.CVC.CC

The particularly important properties of Temiar and Semai prosodic structures are

summarized as follows:

- (49) Characteristic properties of Temiar and Semai prosodic structures:
- a. An onset is obligatory.
 - b. No complex margins are permitted.
 - c. A coda is optional in word-initial/internal positions, but is obligatory in the word-final position.
 - d. Stress falls on the word-final syllable.
 - e. The sesquisyllabic structure with a minor syllable is permitted.
 - f. Minor and major syllables can be concatenated with each other.
 - g. No minor syllable may follow a major syllable within a prosodic word.

Based on these descriptive generalizations, the next section will develop an OT account of Mon-Khmer prosody and reduplication.

5. An OT Analysis of Mon-Khmer Prosody and Reduplication

This section first presents an OT analysis of the regular prosody of Mon-Khmer languages, suggesting a few system-independent constraints that motivate their unmarked prosodic structures (section 5.1). It then provides an OT account of their reduplication, exploring an explanation of it in terms of a meta-constraint on the interaction of prosodic and B-R correspondence constraints (section 5.2).

5.1 Mon-Khmer Prosody

Let us first consider how the characteristic properties of Mon-Khmer prosodic structures listed in (49) are described in OT terms. First of all, assuming that Onset and *Complex are both ranked at the top of the hierarchy, we can account for the observations in (49a) and (49b):

- (50) Onset: The syllable must have an onset.
- (51) *Complex: No more than one segment may link to any syllable margin position.

The observation about a coda in (49c) can be accommodated by ranking Final-C at the top and No-Coda lower than it:

- (52) Final-C: The word-final syllable must have a coda consonant.
- (53) No-Coda: The syllable may not have a coda.

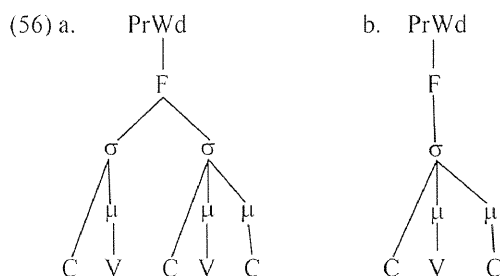
No-Coda favors open syllables, but a closed syllable is required in the word-final position by the more specific and higher-ranking constraint: Final-C.

The basic stress pattern in (49d) is captured by the following two constraints:

(54) Iambic: The foot must be iambic with the form (L \acute{H}) or (\acute{H}).

(55) Align-Ft (Ft, R, PrWd, R): The right edge of every foot must be aligned with the right edge of a prosodic word.

Ranking these constraints at the top of the hierarchy ensures that a prosodic word always has only one foot at the right edge of it, and that the foot takes the shape (L \acute{H}) or (\acute{H}). Given this ranking, together with the assumption that codas are moraic, we can predict that the following prosodic structures are legitimate in Mon-Khmer languages:



This is borne out by the existence of words with an iambic foot like (35) - (37). Note that such words could not be described in usual prosodic terms (i.e., as prosodic words with an iambic foot), *if* codas were not moraic. The consistent stress pattern in Mon-Khmer languages can thus be taken as a piece of evidence suggesting that they *are* moraic.

What about sesquisyllabic structures with a minor syllable? How can we account for the observation in (49e)? One of the crucial constraints is the following:

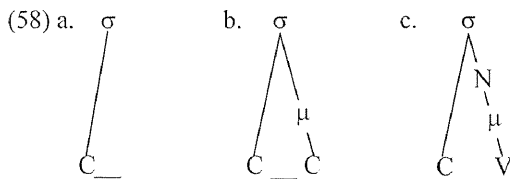
(57) No Unstressed V-Place (NUVP): Vocalic features must be in a stressed position.
(Spaelti 1997: 53)

I assume with Spaelti (1997) that the loss of vowels (or vocalic distinctions) in the unstressed reduplicant is the effect of this constraint. This constraint is not intended to work only for reduplicated forms. Rather, it has broader effects. For instance, it accounts for the well-known observation that vocalic contrasts are often reduced in unstressed positions; e.g., seven vowels are reduced to five in Italian (Vincent 1987), seven to three in Catalan (Mascaró 1978), five to three in Russian (Jones and Ward 1969). It also accommodates similar vowel reduction phenomena (particularly reduction to schwa) that have been observed in many other languages (cf. Selkirk 1977 for French; Hayes 1980/1985 for English and Eastern Cheremis; Cohn 1989

for Indonesian; Kager 1991 for Dutch; Urbanczyk 1995 for Lushootseed; Hayes 1995 for Choctaw, Chickasaw, Ossetic, Cambodian, Araucanian, Macushi, and Cayuga; see Crosswhite 2001 for a comprehensive study of vowel reduction).³⁶ That is, NUVP is responsible not only for the existence of minor syllables in Mon-Khmer languages, but also for the behavior of vowels widely attested in stress-based languages in general.

Given the communicative function of segments (morpheme distinctions) and the speaker's propensity for seeking articulatory economy, this stress-sensitive behavior of vowels makes sense. The vowels in stressed positions, being prominent, are likely to serve the communicative function better than those in stressless positions do. Therefore, the segmental contrasts in the former environment tend to be retained. On the other hand, those in the latter bear less communicative burdens, and are thus more likely subject to reductions under the pressure of economy. Consequently, the constraint in (57) might not have to be stipulated as such. It might rather be explained as the effect of the interaction of the well-known inherently antinomic constraints: (i) the speaker-oriented constraint requiring that articulatory efforts be minimized, and (ii) the hearer-oriented constraint requiring that perceptual distinctions be maximized for communicative purposes (cf. Zipf 1949: 21; Martinet 1960/1964: 167-68, 1962: 139; Carroll and Tanenhaus 1975: 51; Horn 1984: 11, Haiman 1983: 814, Haiman 1985; Du Bois 1985: 358; Krug 2001: 312-14). We shall return to this issue later.

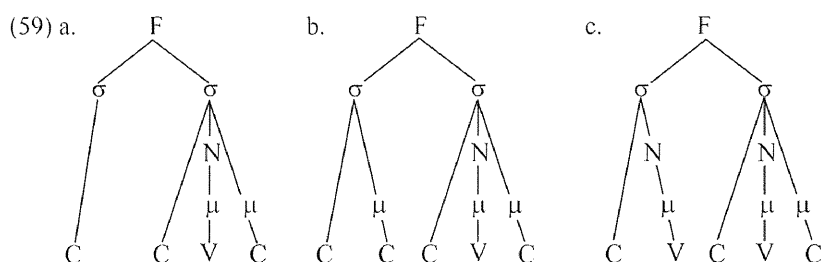
Getting back to the descriptive level, let us consider how minor syllables are represented at the phonological level. I assume, following Shaw (1993, 1996), that they have the prosodic structures in (58a, b), as opposed to the ordinary one in (58c), which is the structure that the major (open) syllable takes:



The crucial difference between the two minor syllables is that the one with a coda in (b) is moraic, whereas the one with no coda in (a) is non-moraic. As indicated earlier, the stress pattern in Mon-Khmer languages evidences the moraic status of the coda.³⁷ This reasoning in turn leads us to the conclusion that the epenthetic vowel in the minor syllable is not moraic. If it were, then it would follow that the minor syllable with a coda (CC) is heavy and does not constitute an iambic foot with the major heavy syllable (CVC) - an unwelcome consequence.

Furthermore, as will be shown shortly, the sesquisyllabic structure is the canonical iambic foot motivated by a domain-general constraint on rhythmic grouping, and the absence/reduction of a vowel in the first syllable (i.e., the very existence of the minor syllable) is explained *as a consequence of its being in the disyllabic iamb*. That is, its form is motivated by *its position in a particular foot* (and by a domain-general constraint), which suggests that the minor syllable with a coda (C_C) is light, and in turn that the epenthetic vowel is not moraic.³⁸ This reasoning is consistent with Piggott's (1995) observation of epenthetic vowels in Mohawk, Iraq Arabic, Selayarese, and Yapese. It shows that certain epenthetic syllables in these languages behave as weightless syllables with respect to some weight-sensitive phenomena like stress assignment, vowel lengthening, and the bimoraicity of the minimal word.

The critical difference between the major and minor syllables is the presence/absence of a nuclear mora. The major syllable in (58c) has a nuclear mora with a vowel, whereas the minor syllables in (58a) and (58b) do not. I take these three syllables as light syllables with different degrees in weight, calling the lightest one in (a) "a non-moraic syllable", the heavier one in (b) "a non-nuclear-moraic syllable", and the even heavier (but still light) one in (c) "a nuclear-moraic (light) syllable". This three-way distinction of light syllables allows us to distinguish the following three kinds of disyllabic feet from one another in terms of weight, thereby opening up a way toward a new dimension of foot hierarchy (cf. Prince 1990):



These are all acceptable in Mon-Khmer languages, which is confirmed by (45), (46) and (37).

Recall here that Mon-Khmer languages have another legitimate foot structure in (56b), and that it, unlike the ones in (59), is monosyllabic. Of these four types of foot structures, I assume that the sesquisyllabic ones in (59a, b) are the unmarked prosodic structures of Mon-Khmer languages. Postponing further discussion until the next section, I here define "unmarked structures" tentatively as ones that have one or more system-independent motivations, which include motivations from language-external systems like perception and motor control. In the case at hand, I assume that the

following system-independent and domain-general constraint plays a crucial role:

(60) Iambic/Trochaic Law:

- a. Elements contrasting in intensity naturally form groupings with initial prominence.
 - b. Elements contrasting in duration naturally form groupings with final prominence.
- (Hayes 1995: 80)

This is a general constraint on rhythmic grouping which motivates us to form certain types of groupings out of a sequence of sounds, and which has been experimentally confirmed with non-linguistic as well as linguistic tasks (cf. Bolton 1894; Woodrow 1909, 1951; Cooper and Meyer 1960; Allen 1975; Bell 1977; Hayes 1995). I assume, following Hayes (1995), that this constraint reflects human perceptual preference, and that (the (b) part of) it exerts strong pressure on the optimal iambic structure. It thus gives system-independent preference to a disyllabic foot with a large durational contrast between the preceding short syllable and the following long syllable, over to one with a smaller or no contrast at all. This constraint, therefore, ensures that the sesquisyllabic foot with a large durational contrast between the two syllables (C(C)-CVC) is a "better iamb" than the disyllabic foot with two major syllables (CV-CVC) and than the monosyllabic foot (CVC), although the latter two are "good enough", i.e., acceptable/legitimate, if not unmarked/canonical.

This analysis is consistent with the following observations made by Hayes (1995) and others. First, there are a great number of languages where the foot with even duration and final prominence (LĹ) is converted to the canonical iamb (LĤ) by lengthening the vowel in the second syllable (e.g., Hixkaryana, Macushi, Surinam Carib, Choctaw, Chickasaw, Menomini, Potawatomi, Cayuga, Onondaga, Seneca, St. Lawrence Island Yupik, Central Alaskan Yupik, Pacific Yupik, Kashaya, Maidu, Sierra Miwok, Yidiñ). Second, there are also languages that increase the durational contrast of the iambic foot by geminating the initial consonant of the following syllable (e.g., Munsee, Unami, Menomini, Seward Peninsula Inupiaq, Central Alaskan Yupik, Pacific Yupik, Southern Paiute). Third, as touched upon earlier, there are a number of languages that make a difference in duration by reducing stressless vowels or their contrasts (see above). Moreover, such lengthening and shortening are much less common in trochaic languages. Finally, it is true that there are a few trochaic languages with lengthening of the stressed vowel (e.g., Chimalapa Zoque, Icelandic, Mohawk), but such (moraic) trochaic languages have shortening of the stressed vowel as well,³⁹ whereas it never occurs in iambic languages. This suggests that it, unlike iambic lengthening, is not motivated by the Iambic/Trochaic Law. In fact, trochaic lengthening tends to be (i) phonetic in the sense that the lengthened vowel is

not as long as the phonologically long vowel and (ii) limited to the main stress syllable, thus suggesting that it is simply a manifestation of stress. Hayes (1995) argues that these observations can all be explained as the effects of the Iambic/Trochaic Law. Due to the rhythmic pressure, iambic languages tend to maintain the durational contrast of the foot by lengthening the stressed syllable and/or shortening the stressless syllable. Consequently, such processes are frequent, robust, and represented structurally at the phonological level. On the other hand, trochaic languages have no motivations to optimize the foot structure by such processes. Therefore, lengthening of the stressed syllable is less common (even if it occurs, it is likely to be a side effect of stress), and even its shortening occurs. Viewing Mon-Khmer sesquisyllabic structure from this perspective, we can take it as a grammaticalized form of the canonical iamb, which is attributed to the above domain-general and system-independent constraint on rhythmic grouping.

Given this claim, one might argue that the sesquisyllabic iamb cannot be the *universally* unmarked prosodic structure, simply because it is not permitted in all the iambic languages, let alone in the trochaic or non-stress-based languages. Obviously, it is not. The unmarked structures of a language were tentatively defined earlier as ones that have one or more system-independent motivations. Such motivations, however, do not always work in perfect harmony with one another. It is often the case that one motivation is in conflict with another (recall the inherent conflict between the speaker-and hearer-oriented constraints touched upon earlier). This is the case for the sesquisyllabic structure, too. The unmarked *foot* structure motivated by the perception system is in conflict with the unmarked *syllable* structure, which is motivated by another language-external system: the motor control system. I assume with MacNeilage and Davis that the canonical syllable structure is CV, and that it is motivated by the universal motor base for speech: a rhythmic open/close alternation of the mandible (cf. MacNeilage and Davis 1990, MacNeilage 1997). This is supported by observations about (i) the child's acquisition order of syllable structures, (ii) the adult's usage of them in natural discourse, and (iii) their typological distribution. First, CV is the type of syllable that overwhelmingly dominates children's vocal output at the early stages of development, including the babbling stage (cf. Vihman 1996). Second, it is used by adult speakers with the highest frequency in natural discourse, even in languages like English that allow very complex syllable structures (cf. Dauer 1983). Third, it is most widely attested in the world's languages: there are almost no languages without it (cf. Blevins 1995; see Breen and Pensalfini 1999 for an arguable exception).

Assuming these perceptual and motoric motivations, I suggest that their conflict is resolved by ranking them in each language, and that the perceptual motivation is favored

in Mon-Khmer languages whereas the motoric motivation is ranked higher in many other languages where CV is omnipresent. That is, language-external motivations provide a set of candidates of system-independently unmarked structures for languages, and each language selects one (or more) of them as its canonical structure(s). To be sure, this analysis does not explain completely why a given language has a certain structure as the canonical one, but does restrict the range of unmarked structures for human languages to a large extent. This system-independent characterization of unmarkedness is one of the main points of this study.

Having discussed motivations for system-independently unmarked structures in general and for the canonical prosodic structures of Mon-Khmer languages in particular, let us next consider how they can be described in OT terms. More specifically, what constraints are required to describe the canonical iambs (CC-C^ˈV and C-C^ˈV), differentiating them from the non-canonical but legitimate ones (CV-C^ˈV and C^ˈV)? First, the following constraint suffices to distinguish mono- and di-syllabic feet:

- (61) Foot Binarity (σ): The foot must be disyllabic.

Given this constraint, the superiority of the disyllabic foot (LH) over the monosyllabic one (H) follows as its effect. Second, as indicated earlier, NUVP accounts for the difference between the sesquisyllabic foot with a minor syllable (C(C)-CVC) and the disyllabic one without it (CV-CVC). This constraint must be ranked high enough to manifest its effect in Mon-Khmer languages, whereas in languages where CV syllables are used with high frequency, it must be outranked by a constraint that masks its effect (e.g., by the one that prohibits syllables with no vocalic nucleus). It is not the case, however, that the effects of these two crucial constraints (Ft-Bin (σ) and NUVP) are always visible in Mon-Khmer languages; for words with no minor syllable like CV-CVC or CVC are also acceptable (e.g. (35) - (37)). This observation suggests that the above two constraints are outranked by the following I-O faithfulness constraints, which militate against deletion and epenthesis:

- (62) Max-IO: Every segment in the input must have a corresponding segment in the output.

- (63) Dep-IO: Every segment in the output must have a corresponding segment in the input.

We have seen the types of words whose syllables are all parsed into a foot. As described above, however, Mon-Khmer languages have words with three or more syllables, where minor and/or major syllables are concatenated with each other (recall

(47) and (49f)). To accommodate such polysyllabic words, we need to assume the following constraint on syllable parsing and to rank it lower than Max-IO:

- (64) Parse- σ : A syllable must be parsed into a foot.

This ranking ensures that Mon-Khmer languages have polysyllabic words, with some syllable(s) unparsed into a foot. For example, a word like (47ai) *k.rn.wa:k* 'frame' is parsed into a foot as follows: *k_.(r_n.wa:k)*.

Let us finally consider the observation in (49g): no minor syllable may follow a major syllable within a prosodic word. None of the prosodic and I-O faithfulness constraints that have been introduced thus far can properly rule out structures like (48). An additional prosodic constraint like the following is required to rule them out:

- (65) *Potential Disyllabic Trochee: A syllable that can bear stress may not precede a syllable that cannot bear stress.

Given this constraint, a string of syllables that *could* be footed as a disyllabic trochee are prohibited altogether even if they *are* footed otherwise, even as a disyllabic iamb. For example, the string of syllables in (48a) could be footed as a disyllabic trochee: [_{PrWd} (CVC.C.)CVC]. It would therefore be ruled out, however it is actually footed, say, as a disyllabic iamb: [_{PrWd} CVC(C.CVC)]. Assuming that (65) stands at the top of the hierarchy, and that major syllables are stress bearing units whereas minor syllables are not, then we can accommodate the observation in (49g).

One may here wonder if the apparently stipulative constraint in (65) can be reformulated as a familiar alignment constraint, but it cannot. For example, if it were formulated as Align (non-moraic σ , L, PrWd, L), then we would fail to rule out illegitimate forms like the following: *CVC.CC.CVC and *C.CV.CC.CVC. If we replaced "non-moraic σ " with "non-nuclear-moraic σ (i.e. minor σ)", then we would fail to rule in legitimate forms like C.CC.CVC and CC.CC.CVC. The constraint in (65) thus needs to be posited as such.

Summarizing, the constraints that are directly responsible for the canonical prosodic structures of Mon-Khmer languages are the following eight: Onset, *Complex, Final-C, Iambic, Align-Ft (Ft, R, PrWd, R), NUVP, Ft-Bin (σ), and Parse- σ . These constraints must all be satisfied simultaneously to yield the sesquisyllabic iambs: (C-C \acute{V} C) and (CC-C \acute{V} C).⁴⁰ It is not the case, however, that they are always satisfied in Mon-Khmer languages. The last three constraints (NUVP, Ft-Bin (σ), and Parse- σ), being outranked by Max/Dep-IO, are sacrificed when their violation is compelled by one or both of the dominating I-O faithfulness constraints. In this case, iambic feet like the following would emerge:

(CV-CVC), (CVC), $C_.(C_C.CVC)$, $C_C.(C_C.CVC)$, $C_.(CV.CVC)$, etc. That is, the above eight constraints are all satisfied when yielding the canonical prosodic structures, but three of them are violated when yielding the acceptable but non-canonical structures. In addition to these constraints, *Potential Disyllabic Trochee is required to rule out strings of syllables where a minor syllable follows a major one. This constraint is never violated and thus must be undominated in Mon-Khmer languages. In contrast to these high-ranking constraints, a prosodic constraint (No-Coda) is often violated in those languages, even when yielding their canonical structures. It must accordingly be ranked low in the hierarchy.

With this constraint ranking in mind, let us see how these prosodic and I-O faithfulness constraints interact with B-R identity constraints to generate the five types of Mon-Khmer reduplication, and how the meta-constraint in (8) gets into the picture.

5.2 Mon-Khmer Reduplication

This section develops an OT analysis of Mon-Khmer reduplication, showing how that meta-constraint works. Answers to the main questions in (6) are given along the way.

Let us begin with the prefixal reduplication of Type 1:

- (66) Mon-Khmer reduplication of Type 1:
 C_1C_r prefixation to the heavy syllable of the base
 Base: a biconsonantal root
 Meaning: active voice and continuative aspect
 $C_1V(:)C_r \rightarrow C_1C_r.C_1V(:)C_r$
- | | | | |
|---------|------|---|----------------------|
| Temiar: | i. | $k\phi:w \rightarrow kw.k\phi:w$ | ‘to be calling’ |
| | ii. | $ca:\phi \rightarrow c\phi.ca:\phi$ | ‘to be eating’ |
| | iii. | $lug \rightarrow lg.lug$ | ‘to be laughing’ |
| Semai: | iv. | $koh \rightarrow kh.koh$ | ‘to be chopping off’ |
| | v. | $ci:p \rightarrow cp.ci:p$ | ‘to be walking’ |
| | vi. | $k\phi:\phi \rightarrow k\phi.k\phi:\phi$ | ‘to be vomiting’ |

The crucial B-R identity constraints required to accommodate this type of reduplication are the following two:

- (67) Max-BR: Every segment of B has a correspondent in R.
 (68) SRole: A segment in R and its correspondent in B must have identical syllable roles.

If we assume SRole and rank Max-BR higher than No-Coda but lower than the other constraints, we can account for this apparently unusual type of reduplication.⁴¹ Let us

see how this ranking works, using several crucial constraints for expository purposes:⁴²

(69) Mon-Khmer Reduplication of Type 1:

[Input: RED, <i>kɔːw</i>]	Onset	*Complex	Iambic	Align-Ft	NUVP	Ft-Bin (σ)	Max-BR	No-Coda	SRole
a. $\text{a.} \text{ (k_w.kɔːw)}$							**	**	
b. $\text{b.} \text{ (_w.kɔːw)}$	*!						***	***	
c. $\text{c.} \text{ (kw_kɔːw)}$		*!					***	*	*
d. $\text{d.} \text{ (kɔːw.k_w)}$			*!				***	***	
e. $\text{e.} \text{ (kɔːw)kɔːw}$				*!	*	*		***	
f. $\text{f.} \text{ (kɔːw)(kɔːw)}$				*!	*	***		***	
g. $\text{g.} \text{ (kɔ.kɔːw)}$					*!		***	*	
h. $\text{h.} \text{ (k_kɔːw)}$							***!	*	
i. $\text{i.} \text{ (w_k.kɔːw)}$							**	**	*!*

Notice first that the candidate in (a) respects all of the first six prosodic constraints (and SRole) at the expense of the lower-ranking constraints Max-BR and No-Coda, whereas the competitors in (b) - (g) all violate one or more of the higher-ranking constraints. Candidates (b) and (c) run afoul of Onset and *Complex, respectively. Candidate (d) breaches another top-ranking constraint: Iambic.⁴³ Candidates (e) and (f) violate Align-Ft, NUVP and Ft-Bin (σ), with (f) incurring one more violation of the binarity condition.⁴⁴ Candidate (g) obeys the two constraints on foot, but violates NUVP. How about candidate (h), then? It, like candidate (a), satisfies all of the first six prosodic constraints, and incurs one less violation of No-Coda than (a) does. However, it incurs one more violation of the higher-ranking constraint Max-BR, thus losing out. Candidate (i) avoids this problem about Max-BR by copying two consonants as (a) does. But it, unlike (a), copies the base onset *k* in the coda position and the base coda *w* in the onset position, which incurs two violations of SRole. Consequently, the sesquisyllabic iamb in (a) correctly wins out as the optimal form.

There are a number of important points to be noted about this reduplication. First of all, this example represents the emergence of the unmarked prosodic structure of Mon-Khmer languages: the disyllabic iamb with a minor syllable (CC-CV̌:C). That is, it can be explained as an instance of the emergence of the unmarked. Second, this unmarked form is created by adding to the base (CV:C) a reduplicant that has an illegitimate prosodic structure in itself (CC). This is a rare type of the emergence of the unmarked; for, as illustrated in section 3, the most common type is such that a prosodically unmarked structure emerges *as a reduplicant*. Third, the shape and position of the reduplicant both fall out of the interaction of independently motivated constraints. Unlike Sloan (1988) and Gafos (1998), neither special templates nor alignment constraints need to be stipulated only for this type of reduplication.⁴⁵ Fourth, the unusual B-R association pattern also follows from the interaction

of usual constraints. Unlike Sloan (1988), no special association rule is required. Fifth, the base form (CV:C) meets the Iambic condition by itself, but the reduplicated disyllabic form is a better iamb than it in terms of Ft-Bin (σ). Thus, this example represents the emergence of the unmarked with respect to Ft-Bin (σ) in the technical sense discussed earlier: the effect of a constraint emerges when a higher-ranking constraint is not applicable. Recall that the effect of Ft-Bin (σ) is not always visible in Mon-Khmer languages, due to the dominating I-O faithfulness constraints, as illustrated in (35) and (36). It emerges, however, in reduplicated forms, because the faithfulness constraints do not apply to the reduplicant, which by definition has no underlying representation. Sixth, the (g) candidate (CV-CV:C) is also better than CV:C in terms of Ft-Bin (σ), but the optimal foot (CC-CV:C) is even better in terms of NUVP. The effect of this constraint, like that of Ft-Bin (σ), is not always visible in the base, since it is outranked by Max/Dep-IO (which is evidenced by (37) and (38)). But it also emerges in reduplication; recall that NUVP plays a decisive role in selecting (a) over (g). It thus turns out that (69) is an example of the emergence of the unmarked with respect to NUVP, too. Seventh, the disyllabic iamb with a non-moraic minor syllable C-CV:C in (h) might be prosodically better than the one with a moraic minor syllable CC-CV:C in (a); recall note 40. In fact, however, the latter is favored over the former in the above competition. This is presumably because C-CV:C does not serve the semantic function of reduplication so well as its rival does. Assuming that the function of reduplication can be satisfied better by copying more segments than less, it follows that CC-CV:C is better than C-CV:C *as a reduplicated form*. This is the effect of Max-BR. Why not always copy all the segments, then? Because this constraint is ranked low enough, crucially, lower than the constraints that are directly responsible for the canonical prosodic structures of Mon-Khmer languages.

Here comes the most important point. Why, then, are these constraints ranked in the way they are in the first place (i.e., in such a way that one of the prosodically unmarked structures emerges whereas marked structures do not)? This is where the meta-constraint in (8) comes in. Given that higher-order constraint, we can answer this 'why' question. The reduplication-specific constraint Max-BR may not be ranked higher than the constraints that are directly responsible for the canonical prosodic structures of the language (Onset, *Complex, Final-C, Iambic, Align-Ft, NUVP, Ft-Bin (σ), and Parse- σ), although it may (and does) dominate other prosodic constraints like No-Coda.

Notice here that the meta-constraint in question does not completely prohibit reduplication-specific constraints from outranking prosodic constraints, and that (69) shows that such a ranking possibility should be permitted. Max-BR must outrank No-Coda; otherwise, the candidate with an open syllable reduplicant in (h) would

incorrectly emerge as the optimal form. Therefore, simply fixing the skeletal emergence -of-the-unmarked ranking in (30) does not work.

Notice also that the meta-constraint in (8) basically prohibits a B-R identity constraint like Max-BR from outranking NUVP and Ft-Bin (σ) in Mon-Khmer languages, despite their being, like No-Coda, not ranked at the top of the hierarchy.⁴⁶ Suppose Max-BR and No-Coda are ranked higher than them. It would yield a disyllabic iamb (CV-CV:C) like (69g) as the optimal form. This is likely to happen in many other languages where CC is not legitimate in the regular prosody, because NUVP and Ft-Bin (σ) must be ranked low in such languages. In languages of the Mon-Khmer type, however, it does not happen (unless a vowel is prespecified for the reduplicant).⁴⁷ This is because the two constraints in question, together with some other constraints, are directly responsible for the canonical prosodic structure, and thus the above meta-constraint disfavors their being outranked by the reduplication-specific constraint Max-BR. It thus follows that the sesquisyllabic iamb CC-CV(:)C emerges as the system-dependent unmarked structure in Mon-Khmer languages, and not in others. This constitutes an answer to the main questions posed in (6): Why is it that Mon-Khmer reduplication takes the form that it does? Why is it that Mon-Khmer languages have such apparently unusual types of reduplication whereas many other languages do not? This reasoning in turn suggests that what structure is unmarked may differ depending on the language, and that the notion of (un)markedness needs to be characterized system-dependently as well as system-independently.

Let us turn to the reduplication of Type 2:

(70) Mon-Khmer Reduplication of Type 2:

C_r infixation immediately to the left of the heavy syllable of the base

Base: a bi-/tri-consonantal root (and the causative affix *r/tr*)

Meaning: active (causative) voice and continuative aspect

$C_1(r/tr).CV(:)C_r \rightarrow C_1(t/r)C_r.CV(:)C_r$

- | | | | | | |
|---------|------|---------|---|-----------|-------------------------|
| Temiar: | i. | s.lɔg | → | sg.lɔg | ‘to be lying down’ |
| | ii. | sr.lɔg | → | s.rg.lɔg | ‘to be laying down’ |
| | iii. | tr.kɔ:w | → | t.rw.kɔ:w | ‘to be calling’ |
| Semai: | iv. | c.ʔu:l | → | cl.ʔu:l | ‘to be choking’ |
| | v. | b.he:ʔ | → | bʔ.he:ʔ | ‘to be being satisfied’ |
| | vi. | c.ləh | → | ch.ləh | ‘to be going down’ |

If we assume the constraint on syllable parsing in (64) and rank it above Max-BR, then this infixal reduplication also follows:

(71) Mon-Khmer Reduplication of Type 2:

[Input: RED, s.lɔg]	Iambic	NUVP	Ft-Bin(σ)	Parse- σ	Max-BR	SRole
a. \varnothing (s_g.lɔg)					***	
b. (l_g)(s_.lɔg)	*!		*		**	
c. (sɔg.lɔg)	*!	*			**	
d. (s_.lɔg)(s_.lɔg)		*!				
e. s_.lɔg(lɔg)		*!	*	**	*	
f. s_(l_g.lɔg)				*!	**	
g. s_g(s_.lɔg)				*!	**	
h. (s_l.lɔg)					***	*!

The optimal candidate in (a) satisfies all the constraints except for Max-BR, whereas the alternatives in (b) - (h) do not. Candidate (b) runs afoul of Iambic and Ft-Bin (σ) as well as Max-BR, because it contains an illegitimate and monosyllabic foot. Candidate (c) has a foot that consists of two heavy syllables, thus breaching Iambic and NUVP. Candidate (d) takes the form of total reduplication, which satisfies Max-BR completely. But it violates NUVP since the vowel in the base is copied together with the flanking consonants and appears in the stressless position. Candidate (e) incurs less violations of Max-BR than (a) does, but does not obey the three mid-ranking constraints (NUVP, Ft-Bin (σ) and Parse- σ), all of which (a) respects. Candidates (f) and (g) meet the conditions on foot forms, segmental contents, and B-R association patterns. Moreover, they incur one less violation of the maximal copying constraint than candidate (a) does. However, they sacrifice the higher-ranking constraint on syllable parsing, whereas the optimal candidate does not.⁴⁸ Candidate (h) avoids a violation of NUVP by copying no offending vowel, but the copied segment has a different syllable role from the corresponding segment in the base, which leads to a violation of SRole. It thus turns out that the candidates in (b) - (h) are all less harmonic than (a) and lose out to it.

Note here that the candidates in (g) and (e) look like reduplicated forms of Types 4 and 5, and that they are both correctly ruled out as the optimal candidate for the reduplication of the type under discussion: Type 2. Note also that the present analysis excludes them without stipulating any templatic and alignment constraints for reduplication (whereas the previous analyses posit such devices and yet fail to accommodate Type 2 while distinguishing it from Types 4 and 5; recall note 45).

Let us move on to the third type:

(72) Mon-Khmer Reduplication of Type 3:

C₁ prefixation with the fixed vowel *a* to the heavy syllable of the base

Base: a biconsonantal root

Meaning: i. active voice and simulfactive aspect

- ii. resultatives
 $C_1 V(:) C_f \rightarrow C_1 a.C_1 V(:) C_f$
- Temiar: i. $k\acute{o}:w \rightarrow ka.k\acute{o}:w$ ‘to call’
 Semai: ii. $r\acute{o}c \rightarrow ra.r\acute{o}c$ ‘to be uprooted (of a tuber)’
 iii. $c\tilde{e}:s \rightarrow ca.c\tilde{e}:s$ ‘to be torn off’

To handle this type of reduplication, a prespecification of the vowel *a* must be made, but no more assumptions are necessary:⁴⁹

(73) Mon-Khmer Reduplication of Type 3:

[Input: $a^{RED}, k\acute{o}:w$]	Iambic	Align-Ft	Max-IO	NUVP	Max-BR
a. $\text{ts}^{\text{RED}}(ka.k\acute{o}:w)$				*	***
b. $(k\acute{o}:w.k\acute{o}:w)$	*!		*	*	
c. $(kaw)(k\acute{o}:w)$		*!		*	**
d. $(k_w.k\acute{o}:w)$			*!		**

The candidate with the prespecified vowel in (a) satisfies the first three constraints while sacrificing NUVP and Max-BR. The total reduplication form in (b) respects the B-R identity constraint, but pays high costs for it: a violation of as many as three higher-ranking constraints: Iambic, Max-IO, and NUVP. If the vowel of the first syllable in (b) is replaced by the prespecified one, and if the resulting syllable constitutes a foot on its own, then Iambic and Max-IO are satisfied, as shown in (c). However, it fatally violates another high-ranking constraint (Align-Ft), because it contains a foot that is not properly right-aligned. The candidate with a minor syllable in (d) spares a violation of Iambic, Align-Ft and NUVP by copying no vowel, but instead incurs a fatal violation of the I-O faithfulness constraint. It turns out, consequently, that the optimal reduplicant of Type 3 takes the shape of CV.

It should be emphasized that this light syllable is system-independently unmarked (i.e., universally unmarked in its own right), and that it can nevertheless appear as a reduplicant in Mon-Khmer languages *iff* the vowel is prespecified. This is because NUVP is ranked relatively high (and the I-O faithfulness constraints even higher) in those languages. Recall that NUVP would otherwise militate against the presence of a vowel in the position that the prefixal reduplicant occupies, and would favor the system-dependent unmarked structure: the sesquisyllabic iamb. The emergence of CV as a reduplicant thus hinges upon the prespecification of the vowel *a*.

Let us next consider the two types of expressive reduplication (Types 4 and 5). I argue that two distinct alignment constraints are required to accommodate them while distinguishing them from each other and from Type 2. Let us see how the OT system works for them in turn, beginning with Type 4:

(74) Mon-Khmer Reduplication of Type 4:

 C_1C_f prefixation to the stem base

Base: a polysyllabic stem

Meaning: expressive (prolongation or continuous repetition in time)

 $C_1(.CC).C\ V(:)C_f \rightarrow C_1C_f.C_1(.CC).CV(:)C_f$


- Temiar:
- i. b.guy \rightarrow **by**.b.guy ‘to waft (smoke)’
 - ii. r.we:g \rightarrow **rg**.r.we:g ‘to stand conspicuously upright’
 - iii. k.rd.lā:d \rightarrow **kd**.k.rd.lā:d ‘curly hair’
- Semai:
- iv. d.ŋɔh \rightarrow **dh**.d.ŋɔh ‘appearance of nodding constantly’
 - v. g.hup \rightarrow **gp**.g.hup ‘irritation on skin’
 - vi. b.ʔəl \rightarrow **bl**.b.ʔəl ‘painful embarrassment’

For this type of expressive reduplication, the following alignment constraint is required:

- (75) Align (Affix₄, R, Stem, L): The right edge of Affix₄ must be aligned with the left edge of the stem, where ‘Affix₄’ is expressive (prolongation or continuous repetition in time).

If we rank this constraint higher than Parse- σ and Max-BR, then the optimal form emerges, as shown in the following tableau:

(76) Mon-Khmer Reduplication of Type 4:⁵⁰

[Input: RED, g.hup]	*PDTrochee	NUVP	Align-Affix ₄	Parse- σ	Max-BR
a.  g_p -(g_.húp)				*	**
b. g_.hup -(g_.húp)	*!	*		**	
c. g_- -(g_.húp)				*	***!
d. (g_ p .húp)			*!		***
e. g_hup-(húp)		*!	*	**	*

Candidate (a) satisfies the first three constraints including Align-Affix, at the expense of the two lower-ranking constraints. All the other candidates fare no better than it, and it is correctly selected for this type of reduplication. Candidate (b) loses out at the first competition, since it breaches the top-ranking constraint: *Potential Disyllabic Trochee. Candidate (c), like candidate (a), obeys the first three constraints by less-than-full copying at the designated edge, but incurs one more violation of Max-BR. It thus yields to (a), perhaps for the same functional reason as (69h) loses against (69a); recall the discussion given above. The candidate in (d) looks like a form for the infixal reduplication of Type 2, but is correctly ruled out as a form for the expressive reduplication in question. Notice here that if there were no alignment constraint, then the candidate of the shape CC.CVC, being prosodically desirable, would beat the candidate in (a). That is, it is the alignment constraint in (75) that plays a crucial role in distinguishing the reduplication of Type 4 from Type 2.

One may here wonder if this analysis requires a violation of the meta-constraint in (8). Yes, it does. Parse- σ is one of the constraints that are directly responsible for the unmarked prosodic structure of Mon-Khmer languages, and it is crucially outranked by the alignment constraint in (75), which should be taken as a reduplication-specific constraint. This is, however, not a serious problem for the present analysis, because the meta-constraint, just like ordinary OT constraints, is posited as a soft constraint that can be violated minimally under duress, and (76) is just an instantiation of it. The motivation that competes with (8) in this case is the one touched upon in the preceding section (though formulated there in slightly different terms): the hearer-oriented constraint requiring that distinct meanings be expressed unambiguously by different forms for purposes of the secure and faithful transmission of information.⁵¹ Recall that the canonical prosodic structure (CC-CVC) is used by other types of reduplication to express a certain meaning: continuative aspect. The use of this form for a rather different meaning "expressive" would therefore run afoul of the above functional constraint. This dilemma is resolved in such a way that the meta-constraint in (8) gives way to the competing constraint. More specifically, it is resolved by violating (8) "minimally", where a minimal violation of a meta-constraint of that type is defined as a minimal promotion of one or more relevant constraints. In the case at hand, the reduplication-specific constraint in (75) is only promoted to a higher position than one of the constraints that are directly responsible for the unmarked prosodic structure of Mon-Khmer languages: Parse- σ .⁵² Recall that this constraint is not top-ranked in the regular prosody, and that there are some cases like (47) where it appears to be violated, independently of reduplication. It thus follows that the promotion of Align-Affix₄ over Parse- σ brings out the emergence not of a markedly deviant structure, but rather of a slightly marked one.⁵³

We have considered how and why (76a) is favored over (76d) as a candidate for the reduplication of Type 4, despite the former being apparently more marked than the latter in terms of prosodic structure. How about candidate (76e), then? That looks like a reduplicated form of Type 5. To be sure, the constraint in (75) suffices to correctly rule it out. But the problem is: as things now stand, a form like (76e) would *never* emerge as a reduplicated word in Mon-Khmer languages. As long as (75) is ranked in the present position and exerts its effect in the way described above, candidate (76e) would always lose to (76a). On the other hand, as pointed out above, if that alignment constraint is abandoned completely, then the unmarked prosodic form in (76d) would always win the competition. Moreover, even if the alignment constraint is generalized in a way that does not specify the right or left edge of the stem to which the reduplicant is attached (and no matter how it is ranked with respect to the other constraints), candidate (76e) would

still be ruled out, because it violates NUVP whereas its rival in (76a) does not. We thus have to keep the specific alignment constraint in (75) for Type 4 and need another one for the arguably suffixal reduplication of Type 5, to which let us turn.

(77) Mon-Khmer Reduplication of Type 5:

CV(:)C_F suffixation to the stem base

Base: a polysyllabic stem

Meaning: expressive (repetition at intervals of time)

$C_1(.CV).CV(:)C_F \rightarrow C_1(.CV).CV(:)C_F.CV(:)C_F$

- | | | | |
|---------|------|----------------------------|--|
| Temiar: | i. | k.ra.hab → k.ra.hab.hab | ‘lip-smacking’ |
| | ii. | k.ra.lɔg → k.ra.lɔg.lɔg | ‘sound of heavy footsteps’ |
| | iii. | c.ra.ʔu:k → c.ra.ʔu:k.ʔu:k | ‘stomach queaziness’ |
| Semai: | iv. | d.yɔ:l → d.yɔ:l.yɔ:l | ‘the appearance of an object floating down a river and getting stuck here and there’ |
| | v. | k.na.rã:c → k.na.rã:c.rã:c | ‘repeated pains of deep wound’ |
| | vi. | g.ra.yul → g.ra.yul.yul | ‘several people shaking something repeatedly’ |

To accommodate this type of reduplication while distinguishing it from Types 2 and 4, the following alignment constraint is required, and it needs to be ranked higher than NUVP and Ft-Bin (σ) as well as Parse- σ :

- (78) Align (Affix₅, L, Stem, R): The left edge of Affix₅ must be aligned with the right edge of the stem, where ‘Affix₅’ is expressive (repetition at intervals of time).

Let us see if the interaction of these constraints correctly gives rise to the optimal reduplicated form of Type 5, with the following tableau:

(79) Mon-Khmer Reduplication of Type 5:

[Input: RED, d_yɔ:l]	Align-Ft	*PD Trochee	Iambic	Align-Affix ₅	NUVP	Ft-Bin	Parse- σ
a. $\text{d_y}\check{\text{ɔ}}\text{:l}-(\text{y}\check{\text{ɔ}}\text{:l})$					*	*	**
b. $(\text{d_y}\check{\text{ɔ}}\text{:l})-(\text{y}\check{\text{ɔ}}\text{:l})$	*!				*	*	
c. $\text{d_y}\check{\text{ɔ}}\text{:l}-(\text{d_y}\check{\text{ɔ}}\text{:l})$		*!			*		**
d. $(\text{d_y}\check{\text{ɔ}}\text{:l})-(\text{d_y}\check{\text{ɔ}}\text{:l})$	*!	*			*		
e. $\text{d_}(\text{y}\check{\text{ɔ}}\text{:l}-\text{d_l})$			*!				*
f. $(\text{d_l_y}\check{\text{ɔ}}\text{:l})$				*!			
g. $\text{d_l}-(\text{d_y}\check{\text{ɔ}}\text{:l})$				*!			*

The candidate in (a) respects all the first four constraints at the expense of the last three. On the other hand, the competitors all fail to satisfy one or two of the dominating constraints.⁵⁴ Candidate (b) violates Align-Ft, candidate (c) *Potential Disyllabic Trochee,

and candidate (d) both, although these losers obey one or two of the lower-ranking constraints, all of which the winner breaches: (b) obeys Parse- σ , (c) Ft-Bin (σ), and (d) both. Candidate (e) satisfies another lower-ranking constraint (NUVP), which all the preceding candidates violate, but it incurs a fatal violation of a higher-ranking constraint: Iambic. The candidates in (f) and (g), unlike the above ones, have legitimate prosodic structures and look like reduplicated forms of Types 2 and 4, respectively. They should thus be rejected as a candidate for the particular type of reduplication under consideration, not for Mon-Khmer reduplication in general. The alignment constraint in (78) exactly serves this function. It correctly rules them out as the candidates for the reduplication of Type 5, but does not commit itself to other types of reduplication. In cases where it is not at stake, as shown earlier, the OT system correctly yields the optimal candidates for Types 2 and 4.

Note here that the meta-constraint in (8) is violated in this type of reduplication, too: the reduplication-specific constraint in (78) is ranked higher than some of the constraints that are directly responsible for the canonical prosodic structures of Mon-Khmer languages: NUVP and Ft-Bin (σ) in addition to Parse- σ . Just as in the case of Type 4, however, this violation is also compelled by the aforementioned functional constraint, and is minimal in the sense defined above. Consequently, here again, what emerges as the optimal reduplicated form is a legitimate structure that falls within the range of Mon-Khmer regular prosody, if not a canonical one.

Another point that should be noted about this analysis is that it explains the shape of the reduplicant for this type of reduplication (as well as for the other four), without stipulating any template. Given the OT system developed above, together with the assumption that Type 5 is suffixal reduplication, then it *follows* that the reduplicant takes the shape of CV(:)C.

We have accommodated the reduplication of Types 4 and 5 while distinguishing them from each other and from Type 2. Interestingly, in fact, the two types of expressive reduplicants can co-occur in Mon-Khmer languages, which is exemplified by (80):

- (80) d.yõ:l → dl.d.ra.yõ:l.yõ:l ‘appearance of several objects floating down in several places’ (*ra* is an infix indicating that “the pattern, the movement, the sensation or the sound occurs in several places” (Diffloth 1976b: 253))

The present analysis can even account for double reduplication forms like this, without any additional assumption:

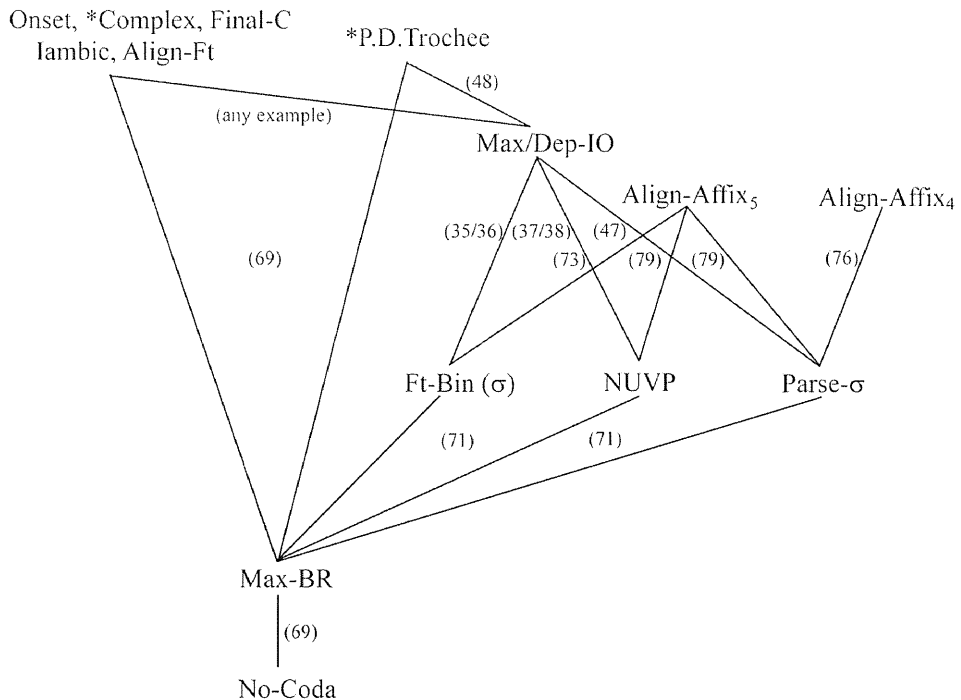
(81) Mon-Khmer Reduplication of Types 4 and 5:^{55, 56}

[Input: RED ₄ , RED ₅ , d.yõ:l]	Align-Affix ₄	Align-Affix ₅	NUVP	Parse-σ
a. d_l.d_.yõ:l-(yõ:l)			*	***
b. d_.yõ:l-(yõ:l)	*!		*	**
c. d_l-(d_.yõ:l)		*!		*

The optimal candidate in (a) satisfies both align-affix constraints by having two reduplicants in the appropriate positions. On the other hand, the failed candidates in (b) and (c) obey only one of them, thus both losing out as the double reduplication form.

This section has demonstrated that the five types of Mon-Khmer reduplication are all derived from the interaction of (i) several constraints on the regular prosody of the language, (ii) familiar I-O and B-R correspondence constraints, and (iii) two types of alignment constraints for expressive reduplication. These constraints and their ranking are summarized as follows:

(82) Constraint Ranking in Temiar and Semai:



We should here pay attention particularly to the ranking of reduplication-specific constraints relative to prosodic ones. Max-BR is ranked lower than all the constraints that are directly responsible for the canonical prosodic structures of Mon-Khmer languages. The two types

of Align-Affix constraints are minimally promoted to a position that is higher than the three mid-ranking constraints that are sometimes violated independently of reduplication: Ft-Bin (σ), NUVP and Parse- σ . There are no cases where reduplication-specific constraints outrank the otherwise undominated constraints. This is explained as the effect of the meta-constraint in (8): given that higher-order constraint, such cases should be highly unlikely, if any. For comparison, consider (76) again, with an unlikely ranking like the following:

(83) Reduplication in Semai' (a hypothetical language):

[Input: RED, <i>g.húp</i>]	Max-BR	*PDTrochee	NUVP	Align-Affix _i	Parse- σ
a. <i>g_p</i> -(<i>g_.húp</i>)	*!*				*
b. *☞ <i>g_.húp</i> -(<i>g_.húp</i>)		*	*		**
c. <i>g_-</i> -(<i>g_.húp</i>)	*!***				*
d. (<i>g_p</i>).húp	*!***			*	
e. <i>g_húp</i> -(<i>húp</i>)	*!		*	*	**

If Max-BR were ranked higher than the other constraints, then the total reduplication form in (b) would be selected as the optimal form. This is, however, contrary to the fact that such a structure with a major syllable preceding a minor syllable is not attested in Mon-Khmer languages. As pointed out earlier, a problem with OT is that it freely allows such an emergence-of-the-marked ranking pattern. If it is complemented with the meta-constraint in (8), however, many impossible/unlikely types of reduplication like this are correctly excluded. That is, that higher-order constraint contributes to properly delimiting the range of possible/likely types of reduplication and in turn of grammars, capturing the ubiquity of the emergence of the unmarked and the strong dependency between a reduplicated form and the canonical prosodic structure of the language: what reduplicated form a given language has depends strongly on what prosodic structure it has.

Here arises a question: Why is it, then, that languages strongly tend to have such prosodic dependency in the first place? When it comes to this fundamental question, the meta-constraint formulated in (8) does not help. And again, as indicated at the very beginning of the paper, strong dependencies hold not only among prosodic structures but also among linguistic entities in general; recall the dependency among phonological segments illustrated in (2). The above question can thus be reformulated in more general terms: Why is it that languages tend to have strong dependencies among their elements? The meta-constraint in (8) is too specific to handle this general question. That higher-order constraint itself needs to be explained in terms of a more general constraint. Let us finally consider this question as a direction for future research.

6. Direction for Future Research: Exploring the Origins of Linguistic Dependency

The strong dependency among elements is one of the most essential properties of language (or of systems in general for that matter); recall Jakobson and Waugh's claim quoted earlier. I suggest that it might be explained as the effect of a constraint requiring that the elements of a language system share maximal information with one another. The information of an entity is defined as a set of features that differentiate it from others, where features may be of any type: e.g., phonological, semantic or symbolic. The degree of information sharing is measured in terms of the ratio of the number of elements to the number of features that characterize them.⁵⁷ Given this constraint, it follows that, the number of features being equal, the more elements a language system has, the more information sharing it has. The number of elements being equal, on the other hand, the degree of information sharing increases in inverse proportion to the number of features that are required to differentiate them from one another. This constraint thus motivates the speaker to carve up the world into segments in a way that maximizes their number with least features, thereby increasing the degree of information sharing among them.⁵⁸ It can therefore be taken as a speaker-oriented constraint. I assume, here again, that it can be violated minimally under compulsion of one or more competing hearer-oriented constraints like the one indicated earlier.

For example, unless otherwise motivated, "the consonantal sound space" is carved up into segments in such a way that they share information with one another maximally. Consonants thus segmented inevitably share information with many other consonants, where information may be formulated in terms of features of the place and manner of articulation. To take (1) and (2) again as examples, assuming three place-of-articulation features (bilabial, alveolar and velar) and two manner-of-articulation features (voice and nasal), then the segment/feature ratio of the consonantal inventory in (1a) is 9/5; that is 1.8. On the other hand, that of the inventory in (1b) is, with six additional place-of-articulation features, 9/11 or only about 0.82. This means that (1a) has much higher degree of information sharing than (1b) does, which makes them appear markedly different in systematicity. As pointed out above, actual languages tend to have a highly (if not completely) systematic paradigm similar to (1a), rather than a random list like (1b), and what consonant a given language has depends strongly on what other consonants it has.⁵⁹ That is, they tend to have a relatively high degree of information sharing. For example, each nasal of Sui in (2a) shares a place-of-articulation feature with two other nasals and a manner-of-articulation feature with three others. Assuming four place-of-articulation features (bilabial, alveolar, palatal and velar) and two manner-of-articulation features (voice and creak), then the segment/feature

ratio of the nasal system of Sui is $12/6$ or 2 , a high degree of information sharing. The affricates of Chipewyan in (2b) also take a three-by-four system with the same degree of information sharing. Alawa has a system of stops with a higher degree of information sharing, which is, with five place-of-articulation features and two manner-of-articulation features, $15/7$ or about 2.14 . It seems to be this multiple information sharing that works as a strong force for attracting consonants to one another, thereby giving rise to a system with well-aligned and strongly interdependent elements. That is, it seems that their systematicity and dependency are attributed to the high degree of information sharing among them, and in turn to the above constraint that is responsible for its emergence.

Essentially the same explanation is likely made for the emergence of a prosodic dependency (or of linguistic dependencies in general for that matter). Once a prosodic unit is established as the canonical structure in the grammar of a language, then it comes to be given system-dependent preference by the above constraint, and it becomes likely to be utilized for enhancing expressive power. For example, once the sesquisyllabic iamb is entrenched in the grammar of a Mon-Khmer language, then it becomes more likely to be exploited for the purpose of reduplication. If this recycling of the well-entrenched structure for a new function takes place, then a dependency inevitably emerges between a reduplicated form and the canonical prosodic structure of the language. Hence the main claim of the present study: what reduplicated form a given language has depends strongly on what prosodic structure it has. This is essentially equivalent to the practical effect of the meta-constraint in (8): reduplication-specific constraints strongly tend to be ranked lower than constraints that are directly responsible for the unmarked prosodic structure(s) of the language; for this ranking yields the prosodic dependency in the above sense. This reasoning in turn suggests that (8) might not have to be stipulated as such, but might rather be derived as a theorem from the more general constraint on information sharing.

Under this analysis, consequently, the apparently unusual types of reduplication in Mon-Khmer languages turn out to make perfect sense. Given their canonical prosodic structure, which is motivated independently of reduplication, adopting the sesquisyllabic form for it is the optimal way for them to enhance expressive power while maximizing information sharing among their elements. Mon-Khmer reduplication thus turns out to be explained as an instance of the emergence of the unmarked in the particular sense that has been discussed throughout the paper: the independently-motivated canonical structure of a language is utilized for a new function and makes its appearance.

Notes

* I would like to thank Norio Yamada for helpful comments on an earlier version of this paper. Any remaining shortcomings are my responsibility.

¹ The columns and rows in the inventories indicate the place and manner of articulation, respectively. Asterisks are meant to show the absence of the phonemes in the positions that they occupy.

² Mon-Khmer and Aslian are (sub)families of Austroasiatic languages, but there is no consensus on their classification and naming: some researchers take them as separate groups, whereas others use the term "Aslian" to refer to a subgroup of Mon-Khmer (cf. Ruhlen 1987: section 4.5). The arguments to be made later do not hinge on this grouping, and this paper adopts the more familiar term "Mon-Khmer" to refer to the language group that Temiar and Semai belong to.

³ The following data of Temiar and Semai are taken from Benjamin (1976) and Diffloth (1976a, b).

⁴ This paper adopts a strictly phonological orthography and does not write arguably epenthesized vowels for several reasons to be presented later (see section 5 for discussion). Most of the previous studies on Mon-Khmer languages also adopt the phonological notation.

⁵ The base is represented to the left of the arrow, and the reduplicated form to the right. C_1 and C_r stand for the first and final consonants of the base, respectively. The reduplicant is highlighted by boldface.

⁶ The two forms (*r* and *tr*) are the allomorphs of the causative affix.

⁷ The notion of "expressive" is hard to define. Benjamin (1976) describes it as follows: "Semantically, they [expressives] serve as a kind of expressive mirror-phrase, summing up in a word or two the 'feelings' that are stereotypically supposed to be aroused in the interlocutors' minds. It is extremely difficult to find satisfactory translation labels for these forms because, even though they are standardized phrases, they are concerned more with connotational than with denotational meanings. They are very common in ordinary conversation, and in stories and song-lyrics they are an essential element of the style." (p. 177) Diffloth (1976b: 251-52) claims that expressive reduplication morphemes connote "prolongation or continuous repetition in time" or "repetition at intervals of time".

⁸ It is not very clear whether the heavy syllable reduplicant is suffixed or infix, but I tentatively assume that it is a suffix for reasons to be presented in section 5.

⁹ Semai is not so consistent as Temiar in terms of vowel epenthesis. For example, *a* is inserted before *h* and *ʔ*, and thus the phonetic realization of (1aiv) is *kahkoh*.

¹⁰ Moreover, it will be shown later that the problem of copying the vowel in the base should be attributed to the prosodic markedness of a reduplicated form as a whole rather than to the segmental markedness of the vowel itself (see section 5).

¹¹ This sense of the emergence of the unmarked is not identical with the technical and narrower sense

of the term in OT, which will be discussed in section 3.

¹² Some proposals have been made to prevent the proliferation of possible grammars (e.g., fixing a certain constraint ranking universally). But it will be shown later that they do not help to explain the empirical phenomena to be taken up in this paper (see section 3).

¹³ Kammu reduplication, which Svantesson (1983) and Takeda (1997) take up, is not discussed in this paper, although a few brief notes are made in comparison to Temiar and Semai.

¹⁴ Apparently based on this description by Sloan (and on the description by Gafos 1998), Raimy (2000: 146) makes a misleading statement to the effect that Temiar and Semai are different in terms of the availability of reduplication.

¹⁵ Here again, Raimy (*ibid.*) apparently inherits the same problem, and makes an inaccurate claim that reduplicated forms like (9) as well as ones like (10) indicate continuative aspect.

¹⁶ The parentheses indicate optionality; thus, (11) represents two structures: [CC-CVC]_{PrWd} and [C-CVC]_{PrWd}.

¹⁷ Not all studies framed in terms of OT explore a non-templatic approach: see Downing (2000) for arguments against obviating templates.

¹⁸ Gafos (1998) does not deal with Semai reduplication.

¹⁹ It will be shown later that whatever segment is copied in the coda position, the resulting CVC structure is less harmonic than CV with respect to certain constraints on the canonical prosodic structure of Mon-Khmer languages, and that the prosodic constraints play a crucial role in the explanation of Mon-Khmer reduplication (see section 5).

²⁰ He does not show explicitly how this works with a tableau. The tableau in (21) is one that I have made up.

²¹ Even if stress falls on the infixed reduplicant, candidate (e) is ruled out by Markedness.

²² Incidentally, another previous OT analysis of Mon-Khmer reduplication, Hendricks (2001), deals only with Type 4 (C_iC_r prefixation to the stem) in Semai, leaving the other four types of reduplication unexplained.

²³ Notice that if the base were not a morphological stem but a stressed syllable, then a reduplicated form like the following would incorrectly emerge: g.*hp.hup*.

²⁴ The following examples are, of course, not meant to be exhaustive; for other examples of reduplication, see the references to be cited below, particularly McCarthy and Prince's series of papers on prosodic morphology (McCarthy and Prince 1986/1996, *et seq.*).

²⁵ As has been pointed out in the literature, a syllabically unmarked structure is not necessarily unmarked in other respects (e.g., with respect to the metrical constraint on the "minimal word"), which will be illustrated shortly.

²⁶ The word-final consonant of this example might be extrasyllabic. If so, then this word would turn out to be an example of total reduplication. A similar note will apply to some of the other examples in (24) - (27). For example, if the word-final consonant in (24di) turns out to be extrasyllabic, then that example would be categorized into the class of the emergence of CV from CVC rather than from CVCC, although it still represents the emergence of the unmarked in certain respects (e.g., with respect to No-Coda).

²⁷ This type of optimization is also observed in reduplicated forms with misaligned B-R associations like the following, where CVC emerges from CV; see Jeanne (1982) and Crowhurst (2004) for this type of reduplication:

- (i) laho → **la**-l.ho 'bucket sg./pl.' [Hopi]
- (ii) ca.qap.ta → **ca**-c.qap.ta 'dish sg./pl.'
- (iii) pi.ta.nak.ci → **pi**-p.ta.nak.ci 'hat sg./pl.' (Jeanne 1982, Crowhurst 2004)

²⁸ This paper does not discuss the emergence of a segmentally unmarked or less marked structure in detail, but it is also widely attested (cf. Prince and Smolensky 1993/2004; McCarthy and Prince 1994, 1995; Itô et al. 1995; Alderete et al. 1999; Spaelti 1997; Lombardi 2001).

²⁹ To restrict the possibilities of constraint rankings, some researchers have proposed meta-constraints that ensure a universal nonpermutable constraint hierarchy. Most of them, however, concern harmonic alignment of prominence scales like sonority (cf. Prince and Smolensky 1993/2004: chapter 8; see de Lacy 2004 for arguments against their fixed constraint ranking). Since such prominence scales represent system-*independent* markedness, they do not help to deal with the emergence of system-*dependently* unmarked structures (recall that what structure emerges as the unmarked may differ depending on the language). Even meta-constraints concerning the I/B/R correspondence like "I-B Faithfulness >> I-R Faithfulness" do not exclude the ranking pattern in (31) (cf. Prince and Smolensky 1993/2004: chapter 9; McCarthy and Prince 1995: section 6).

³⁰ It seems to be this characteristic property of Mon-Khmer reduplication that has motivated Hendricks (2001) to abandon a prosodic templatic approach and to instead pursue an alternative, which relies crucially on morphological alignment/anchoring constraints. I agree with him that language-particular prosodic templates are not required to accommodate Mon-Khmer reduplication, but disagree in that the apparently unusual shape of Mon-Khmer reduplicant is attributed primarily to morphological constraints. As will be demonstrated later, even the bare-consonantal augment is prosodically motivated.

³¹ The parentheses indicate optionality, the Kleene stars repeatability, and the underline an epenthesis position. Thus, the diagram in (34) is meant to represent several types of prosodic structures including sesquisyllabic ones, rather than just a single structure. It is assumed on grounds to be presented below

that the coda is moraic whereas the epenthesized vowel is not.

³² There are Mon-Khmer languages (e.g., Kammu, Cua, Chrau, etc.) where an open syllable is permitted in the word-final position, although the syllable in that position must be heavy.

³³ Complex onsets are acceptable in Kammu. Thus it has minimal pairs like the following:

(i) klo:k 'bamboo bowl'

(ii) k.lo:k 'slit drum'

³⁴ The first syllables in these examples happen to be all CVN (nasal), but it is not clear from the available data whether this is just an accident or not (see (47) for similar examples). If the nasal turns out to be non-moraic, then it follows that the examples in (38), like those in (37), represent a light-heavy syllable sequence, which is a prosodically natural consequence (see section 5.1 for discussion of this matter). Further investigations need to be made to decide the moraic status of the nasal coda. I leave this problem for future research.

³⁵ This is one of the grounds for assuming that the epenthesized vowel is not moraic. See the next section for discussion of this matter.

³⁶ Tonal contrasts are also reduced in unstressed positions (cf. Yip 1980/1990 for Mandarin).

³⁷ It is also corroborated by the observation that minor syllables bear a tone in Kammu only when they have a coda (cf. Svantesson 1983).

³⁸ Note that this analysis implies that the minor syllable with no coda (C₋), being non-moraic, is lighter than the minor syllable with a coda.

³⁹ Hayes suggests that such shortening is also functionally motivated in the sense that it makes a string of syllables with even duration and allows a maximal parse of syllables into perfect moraic trochees.

⁴⁰ Given the Iambic/Trochaic Law and the three-way distinction of light syllables proposed above, it follows that C-CVC is a better iamb than CC-CVC, since the minor syllable of the former is lighter than that of the latter.

⁴¹ SRole is crucial for the selection of the optimal candidate, but its ranking with respect to the other constraints is not.

⁴² SRole is tentatively put at the end of the hierarchy in the following tableau, but, as stated just above, it is neutral in terms of ranking.

⁴³ The following anchoring constraint may also play a role in ruling out candidate (d), which asserts system-independent preference to prefixal (rather than suffixal/infixed) reduplication:

(i) Left-Anchor (Base, Reduplicant): The left edge of the reduplicant corresponds to the left edge of the base. (cf. Nelson 2002)

⁴⁴ On the other hand, to be precise, (e) incurs one more violation of the constraint on syllable parsing in (64), because the second syllable is not footed.

⁴⁵ This does not mean, however, that no alignment constraints are needed for *any* type of reduplication in Mon-Khmer languages. As we see later, reduplication-specific alignment constraints are not needed for the reduplication of Types 1-3, but *are* needed for the two types of expressive reduplication (Types 4 and 5). As pointed out in section 2, Gafos's analysis requires a reduplication-specific alignment constraint for Types 1-3 (whereas the alternative analysis being developed here does not), and it does not accommodate Types 4 and 5 (whereas the present analysis does).

⁴⁶ Recall that the meta-constraint is always qualified by "unless otherwise motivated"; hence the hedge "basically". This codicil is required to accommodate cases where a non-canonical structure emerges in reduplication, which will be illustrated later.

⁴⁷ This qualification, too, is a descriptive necessity, which will also be made clear later.

⁴⁸ This effect might be obtained by a constraint on minimizing structure (*Struc) instead of Parse- σ , because (f) and (g) create one more syllable than (a) does.

⁴⁹ Even this vowel prespecification may be explained as an instance of the emergence of the unmarked at the segmental level, perhaps along the line of Alderete et al. (1999).

⁵⁰ Stem boundaries are indicated by hyphens.

⁵¹ The conflict between this functional constraint and the meta-constraint in (8) might be taken as an instance of the familiar conflict between the speaker-and hearer-oriented constraints indicated earlier; for, as will be shown in the next section, (8) might be derived from a more general constraint on language, which could reasonably be interpreted as a speaker-oriented constraint.

⁵² This could be reformulated alternatively as a minimal *demotion* of a prosodic constraint below a reduplication-specific one. Still, I suggest that the above formulation is more appropriate, because (8) is a constraint that takes independently-motivated prosodic constraints as given, and requires reduplication-specific constraints to be ranked in a certain position relative to them, and not vice versa. This suggestion may have some nontrivial implications for the learnability of OT grammars (cf. Tesar and Smolensky 2000).

⁵³ This line of explanation can likely apply to the emergence of a slightly marked structure in Mangap-Mbula reduplication, too.

⁵⁴ Since stress usually falls only on the word-final position in Mon-Khmer languages, I here assume tentatively that it does when the heavy syllable is reduplicated, too. If stress turns out to fall on both the base and the reduplicant, then they would each constitute a prosodic word, and the following structure would be optimal where the base as well as the reduplicant is footed: $(d_y\acute{\sigma}:l)-(y\acute{\sigma}:l)$. Unfortunately, there are no decisive data available about this problem.

⁵⁵ As we saw earlier, Align-Affix₄ should be ranked above Parse- σ , and Align-Affix₅ above NUVP and Ft-Bin (σ) as well as Parse- σ . But there is no empirical reason to rank Align-Affix₄ with respect to

constraints other than Parse- σ . In (81), it is tentatively put in the same position as Align-Affix₅.

⁵⁶ For the sake of simplicity, the irrelevant infix *ra* is put aside in the tableau.

⁵⁷ This idea is essentially the same as Clements's (2003) "feature economy", though different in the range of application: Clements's economy principle is meant to govern the organization of a segmental (particularly, consonantal) system, whereas the present constraint is meant to apply not only to a system of individual segments but also to a higher-order combinatorial system.

⁵⁸ It also motivates the speaker to combine linguistic entities into larger chunks in such a way that higher-order information sharing emerges maximally.

⁵⁹ The qualification in the parentheses "not completely" is a descriptive necessity. The non-systematic aspect of a consonantal system may be due partly to the aforementioned hearer-oriented constraint: the one which requires that perceptual distinctions be maximized for communicative purposes. It may also be due to a system-independent constraint on learning phonological segments. The ease of learning a given segment is determined not only by a system-dependent factor like information sharing, but also by a system-independent one. For example, it is expected for biomechanical reasons that, *other things being equal*, simple bilabial stops like /b/ are easier to articulate and learn than, say, prenasalized stops like /^mb/ for any child speaking any language; see Lindblom et al. (1993) for discussion of this matter. The interaction of such system-independent as well as system-dependent constraints should be taken into account for a proper description and explanation of the systematicity of language.

References

- Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Suzanne Urbanczyk (1999) Reduplication with fixed segmentation. *Linguistic Inquiry* 30: 327-364.
- Allen, George D. (1975) Speech rhythm: Its relation to performance universals and articulatory timing. *Journal of Phonetics* 3: 75-86.
- Austin, Peter (1981) *A grammar of Diyari, South Australia*. Cambridge, UK: Cambridge University Press.
- Beckman, Jill N., Laura W. Dickey, and Suzanne Urbanczyk (eds.) (1995) *University of Massachusetts occasional papers 18: Papers in Optimality Theory*. University of Massachusetts, Amherst: Graduate Linguistic Student Association.
- Bell, Alan (1977) Accent placement and perception of prominence in rhythmic structures. In: Larry Hyman (ed.) *Studies in stress and accent*, 1-13, University of Southern California, Los Angeles.
- Benjamin, Geoffrey (1976) An outline of Temiar grammar. In: Jenner et al. (eds.), 129-187.

- Blevins, Juliette (1995) The syllable in phonological theory. In: John A. Goldsmith (ed.) *The handbook of phonological theory*, 206-244. Oxford, UK: Blackwell.
- Bolton, Thaddeus L. (1894) Rhythm. *American Journal of Psychology* 6: 145-238.
- Breen, Gavian, and Rob Pensalfini (1999) Arrernte: A language with no syllable onsets. *Linguistic Inquiry* 30: 1-25.
- Broselow, Ellen I. (1983) Salish double reduplications: Subjacency in morphology. *Natural Language and linguistic Theory* 1: 317-346.
- Broselow, Ellen I., and John J. McCarthy (1983) A theory of internal reduplication. *The Linguistic Review* 3: 25-88.
- Buckley, Eugene (1994) *Theoretical aspects of Kashaya phonology and morphology*. Stanford, CA: Center for the Study of Language and Information.
- Cabrera, Neonette C, et al. (1965) *Beginning Tagalog*. Berkeley: University of California Press.
- Carroll, John. M., and Michael. K. Tanenhaus (1975) Prolegomena to a functional theory of word formation. *Papers from the parasession on functionalism, Chicago Linguistic Society*: 47-62.
- Churchward, Maxwell C. (1978) *Rotuman grammar and dictionary: Comprising Rotuman phonetics and grammar and a Rotuman-English dictionary*. New York: AMS Press.
- Clements, G. N. (2003) Feature economy in sound systems. *Phonology* 20: 287-333.
- Cohn, Abigail (1989) Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7: 167-216.
- Cooper, Grosvenor, and Leonard Meyer (1960) *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Crosswhite, Katherine (2001) *Vowel reduction in Optimality Theory*. New York: Routledge.
- Crowhurst, Megan (2004) Mora alignment. *Natural Language and Linguistic Theory* 22: 127-177.
- Dauer, Rebecca M. (1983) Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.
- de Lacy, Paul (2004) Markedness conflation in Optimality Theory. *Phonology* 21: 145-199.
- de Saussure, Ferdinand (1916) *Cours de linguistique générale*. Paris: Payot. Translated by Roy Harris (1983) *Course in general linguistics*. London: Duckworth.
- Diffloth, Gerard (1976a) Minor-syllable vocalism in Sinoic languages. In: Jenner et al. (eds.), 229-247.
- Diffloth, Gerard (1976b) Expressives in Semai. In: Jenner et al. (eds.), 249-264.
- Dixon, R. M. W. (1977) *A grammar of Yidiñ*. Cambridge, UK: Cambridge University Press.

- Dixon, R. M. W. (1988) *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Downing, Laura J. (2000) Morphological and prosodic constraints on Kinande verbal reduplication. *Phonology* 17: 1-38.
- Du Bois, John W. (1985) Competing motivations. In: John Haiman (ed.) *Iconicity in syntax: Proceedings of a symposium on iconicity in syntax*, 343-365. Amsterdam: John Benjamins.
- Fitzgerald, Colleen M. (2000) Vowel hiatus and faithfulness in Tohono O'odham reduplication. *Linguistic Inquiry* 31: 713-722.
- Gafos, Diamandis (1998) Eliminating long-distance consonantal spreading. *Natural Language and Linguistic Theory* 16: 223-278.
- Haeberlin, Herman K. (1918) Types of reduplication in the Salish dialects. *International Journal of American Linguistics* 1: 154-174.
- Haiman, John (1983) Iconic and economic motivations. *Language* 59: 781-819.
- Haiman, John (1985) *Natural syntax: Iconicity and erosion*. Cambridge, UK: Cambridge University Press.
- Harrison, Sheldon P. (1976) *Mokilese reference grammar*. Honolulu: The University Press of Hawaii.
- Hayes, Bruce (1980) Metrical theory of stress rules. Doctoral dissertation, MIT, published from Garland, 1985.
- Hayes, Bruce (1995) *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
- Hayes, Bruce, and May Abad (1989) Reduplication and syllabification in Ilokano. *Lingua* 77: 331-374.
- Hendricks, Sean (2001) Bare-consonant reduplication without prosodic templates: Expressive reduplication in Semai. *Journal of East Asian Linguistics* 10: 287-306.
- Hess, Thomas M. (1966) Snohoomism chameleon morphology. *International Journal of American Linguistics* 32, 350-356.
- Horn, Laurence R. (1984) Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In: Deborah Schiffrin (ed.) *Meaning, form, and use in context: Linguistic applications*, 11-42. Washington, D.C.: Georgetown University Press.
- Itô, Junko, Armin Mester, and Jaye Padgett (1995) Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26: 571-613.
- Jakobson, Roman (1962) *Phonological studies*. s'-Gravenhage: Mouton.
- Jakobson, Roman, and Linda R. Waugh (1979) *The Sound shape of language*. Bloomington: Indiana University Press.

- Jeanne, LaVerne M. (1982) Some phonological rules in Hopi. *International Journal of American Linguistics* 48: 245-270.
- Jenner, Philip N., Laurence C. Thompson, and Stanley Starosta (eds.) (1976) *Austroasiatic studies*, Part I. Honolulu, Hawaii: The University Press of Hawaii.
- Johnston, Raymond L. (1980) *Nakanai of New Britain: The grammar of an oceanic language*. Canberra: The Australian National University.
- Jones, Daniel, and Dennis Ward (1969) *The phonetics of Russian*. Cambridge, UK: Cambridge University Press.
- Kager, Rene (1991) Dutch schwa in moraic phonology. *Papers from the 26th regional meeting of the Chicago Linguistic Society*, Part II, 241-255.
- Kennedy, Robert (2002) A stress-based approach to Ponapean reduplication. *Proceedings of the 21st West Coast Conference on Formal Linguistics*: 222-235.
- Krug, Manfred G. (2001) Frequency, iconicity, categorization: Evidence from emerging modals. In: Joan Bybee, and Paul Hopper (eds.) *Frequency and the emergence of linguistic structure*, 309-335, Amsterdam: John Benjamins.
- Laszlo, Ervin (1972) *Introduction to systems philosophy: Toward a new paradigm of contemporary thought*. New York: Gordon and Breach.
- Lichtenberk, Frantisek (1983) *A grammar of Manam*. Honolulu: University of Hawaii Press.
- Lindblom, Björn (1992) Phonological units as adaptive emergents of lexical development. In: Charles A. Ferguson, Lise Menn, Carol Stoel-Gammon (eds.) *Phonological development: Models, research, implications*, 131-163. Timonium, MD: York Press.
- Lindblom, Björn, Diana Krull, and Johan Stark (1993) Phonetic systems and phonological development. In: Benedicte de Boysson-Bardies, Scania de Schonen, Peter W. Jusczyk, Peter F. MacNeilage, and John Morton (eds.) *Developmental neurocognition: Speech and face processing in the first year of life*, 399-409. Dordrecht, Holland: Kluwer.
- Lombardi, Linda (2001) *Segmental phonology in Optimality Theory: Constraints and representations*. Cambridge, UK: Cambridge University Press.
- MacNeilage, Peter F. (1997) Acquisition of speech. In: William J. Hardcastle, and John Laver (eds.) *The handbook of phonetic sciences*, 301-332. Oxford, UK: Blackwell.
- MacNeilage, Peter F., and Barbara L. Davis (1990) Acquisition of speech production: Frames, then content. In: Marc Jeannerod (ed.), *Attention and performance XIII: motor representation and control*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Marantz, Alec (1982) Re reduplication. *Linguistic Inquiry* 13: 435-482.
- Martinet, André (1960) *Eléments de linguistique générale*. Translated by Elizabeth Palmer (1964)

- Elements of general linguistics*. London: Faber and Faber.
- Martinet, Andre (1962) *A functional view of language*. Oxford: Clarendon Press.
- Mascaró, Joan (1978) *Catalan phonology and the phonological cycle*. Bloomington, Indiana: The Indiana University Linguistics Club.
- McCarthy, John J. (1982) Prosodic templates, morphemic templates, and morphemic tiers. In: Harry van der Hulst, and Norval Smith (eds.) *The structure of phonological representations*, Part I, 191-223. Dordrecht, Holland: Foris.
- McCarthy, John J. (2002) *A thematic guide to Optimality Theory*. Cambridge, UK: Cambridge University Press.
- McCarthy, John J., and Alan S. Prince (1986) Prosodic morphology. Manuscript. University of Massachusetts, Amherst and Rutgers University. Issued as Center for cognitive science technical report #32, Rutgers University, 1996.
- McCarthy, John J., and Alan S. Prince (1993a) Prosodic morphology I. Manuscript. University of Massachusetts, Amherst and Rutgers University.
- McCarthy, John J., and Alan S. Prince (1993b) Generalized alignment. In: Geert Booji, and Jaap van Marle (eds.) *Yearbook of morphology 1993*, 79-153. Dordrecht, Holland: Kluwer.
- McCarthy, John J., and Alan S. Prince (1994) The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistic Society* 24: 333-379.
- McCarthy, John J. and Alan S. Prince (1995) Faithfulness and reduplicative identity. In: Beckman, et al. (eds.), 249-384.
- McCarthy, John J., and Alan S. Prince (1999) Faithfulness and identity in prosodic morphology. In: Rene Kager, Harry van der Hulst, and Wim Zonneveld (eds.) *The prosody-morphology interface*, 218-309. Cambridge, UK: Cambridge University Press.
- Miller, John D., and Carolyn P. Miller (1995) Notes on phonology and orthography in several Katuic Mon-Khmer groups in northeast Thailand. *Mon-Khmer Studies* 24: 27-51.
- Miner, Kenneth (1979) Dorsey's law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics* 45: 25-33.
- Nater, H.F. (1984) *The Bella Coola languages*. Ottawa: National Museums of Canada.
- Nelson, Nicole (2002) The prefixing preference in reduplication. *Proceedings of the 21st West Coast Conference on Formal Linguistics*: 320-333.
- Newman, Stanley (1971) Bella Coola reduplication. *International Journal of American Linguistics* 37: 34-38.
- Nivens, Richard (1993) Reduplication in four dialects of West Tarangan. *Oceanic Linguistics*

- 32: 353-388.
- Piggott, Glyne L. (1995) Epenthesis and syllable weight. *Natural Language and Linguistic Theory* 13: 283-326.
- Postal, Paul M. (1968) *Aspects of phonological theory*. New York: Harper and Row.
- Prince, Alan (1990) Quantitative consequences of rhythmic organization. *Parasession on the syllable in phonetics and phonology*, *Chicago Linguistic Society*: 355-398.
- Prince, Alan, and Paul Smolensky (1993) *Optimality Theory: Constraint interaction in generative grammar*. Center for cognitive science technical report #2, Rutgers University. Published from Blackwell, 2004.
- Raimy, Eric (2000) *The phonology and morphology of reduplication*. Berlin: Mouton de Gruyter.
- Rehg, Kenneth L., and Damian G. Sohl (1981) *Ponapean reference grammar*. Honolulu: The University Press of Hawaii.
- Ruhlen, Merritt (1987) *A guide to the world's languages*. Volume 1: *Classification*. Stanford. CA: Stanford University Press.
- Russell, Kevin (1997) Optimality Theory and morphology. In: Diana Archangeli, and Terence D. Langendoen (eds.) *Optimality Theory: An overview*, 102-133. Oxford, UK: Blackwell.
- Sapir, Edward (1921) *Language: An introduction to the study of speech*. San Diego: Harcourt Brace.
- Selkirk, Patricia A. (1977) The French foot: on the status of 'MUTE' *e*. *Studies in French Linguistics* 1: 141-150.
- Shaw, Patricia A. (1993) The prosodic constituency of minor syllables. *Proceedings of the 12th West Coast Conference on Formal Linguistics*: 117-132.
- Shaw, Patricia A. (1996) The non-nuclear status of syllabic obstruents in Berber. Talk given at the 70th Annual Meeting of the Linguistic Society of America. San Diego, CA.
- Sloan, Kelly (1988) Bare-consonant reduplication: Implications for a prosodic theory of reduplication. *Proceedings of the 7th West Coast Conference on Formal Linguistics*, 310-330.
- Spaelti, Philip (1997) Dimensions of variation in multi-pattern reduplication. Unpublished doctoral dissertation, University of California, Santa Cruz.
- Spencer, Andrew (1991) *Morphological theory*. Oxford, UK: Blackwell.
- Steriade, Donca (1988) Reduplication and syllable transfer in Sanskrit and elsewhere. *Phonology* 5: 73-155.
- Stonham, John T. (1994) *Combinatorial morphology*. Amsterdam: John Benjamins.
- Svantesson, Jan-Olof (1983) *Kammu phonology and morphology*. Malmo, Sweden: CWK

Gleerup.

- Takeda, Kazue (1997) Causative formation and single consonant reduplication in Kammu. Manuscript. University of California, Irvine.
- Tarpen, Marie-Lucie (1983) Morphophonemics of Nisgha plural formation: A step towards Proto-Tsimshian reconstruction. *Kansas Working Papers in Linguistics* 8: 123-214.
- Tesar, Bruce, and Paul Smolensky (2000) *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.
- Thomas, David (1992) On sesquisyllabic structure. *Mon-Khmer Studies* 21: 206-210.
- Topping, Donald M. (1980) *Spoken Chamorro*. Honolulu: The University Press of Hawaii.
- Trubetzkoy, N. S. (1939) *Grundzüge der Phonologie*. Translated by Christiane A. M. Baltaxe (1969) *Principles of phonology*. Berkeley, CA: University of California Press.
- Urbanczyk, Suzanne (1995) Double reduplication in parallel. In: Beckman et al. (eds.), 499-531.
- Urbanczyk, Suzanne (1996) Morphological templates in reduplication. *Proceedings of North East Linguistics Society* 26: 425-440.
- Vihman, Marilyn M. (1996) *Phonological development: The origins of language in the child*. Oxford, UK: Blackwell.
- Vincent, Nigel (1987) Italian. In: Bernard Comrie (ed.) *The world's major languages*, 279-302. Oxford, UK: Oxford University Press.
- von Bertalanffy, Ludwig (1968) *General system theory: Foundations, development, applications*. New York: George Braziller.
- Walker, Rachel (2000) Nasal reduplication in Mbe affixation. *Phonology* 17: 65-115.
- Woodrow, Herbert (1909) A quantitative study of rhythm. *Archives of Psychology* 14: 1-66.
- Woodrow, Herbert (1951) Time perception. In: Stanley S. Stevens (ed.) *Handbook of experimental psychology*, 1234-1236. New York: Wiley.
- Yip, Moira J. W. (1980) The tonal phonology of Chinese. Doctoral dissertation, MIT, published from Garland, 1990.
- Zipf, George K. (1949) *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley Press.